

11/11/99

Customer No. 20350

TOWNSEND and TOWNSEND and CREW LLP
Two Embarcadero Center, 8th Floor
San Francisco, California 94111-3834
(415) 576-0200

Attorney Docket No. 018941-000411US

"Express Mail" Label No. EL378169002US

Date of Deposit: November 11, 1999

ASSISTANT COMMISSIONER FOR PATENTS
BOX PATENT APPLICATION
Washington, D.C. 20231

I hereby certify that this is being deposited with the United States
Postal Service "Express Mail Post Office to Addressee" service
under 37 CFR 1.10 on the date indicated above, addressed to:

Sir:

Transmitted herewith for filing under 37 CFR 1.53(b) is the

- ☒ patent application of
☐ continuation patent application of
☐ divisional patent application of
☐ continuation-in-part patent application of

Assistant Commissioner for Patents
Washington, D.C. 20231

By:

Inventor(s)/Applicant Identifier: Richard STEPHENS, Wayne MITCHELL, Sue KALMAN

For: CHLAMYDIA PNEUMONIAE GENOME SEQUENCE

- ☒ This application claims priority from each of the following Application Nos./filing dates:
60/128,606, filed April 8, 1999 and 60/108,279, filed November 12, 1999

the disclosure(s) of which is (are) incorporated by reference.

- ☐ Please amend this application by adding the following before the first sentence: "This application is a ☐ continuation ☐ continuation-in-part of and claims the benefit of U.S. Application No. 60/_____, filed _____, the disclosure of which is incorporated by reference."

Enclosed are:

- ☒ 123 page(s) of specification
☒ 2 page(s) of claims
☒ 1 page of Abstract
☐ sheet(s) of ☐ formal ☐ informal drawing(s).

An assignment of the invention to _____

A ☐ signed ☐ unsigned Declaration & Power of AttorneyA ☐ signed ☐ unsigned Declaration.

A Power of Attorney.

A verified statement to establish small entity status under 37 CFR 1.9 and 37 CFR 1.27 ☐ is enclosed ☐ was filed in the
prior application and small entity status is still proper and desired.

A certified copy of a _____ application.

Information Disclosure Statement under 37 CFR 1.97.

A petition to extend time to respond in the parent application.

Notification of change of ☐ power of attorney ☐ correspondence address filed in prior application.

- ☒ Sequence listing of 203 pages.

In view of the Unsigned Declaration as filed with this application and pursuant to 37 CFR §1.53(f),
Applicant requests deferral of the filing fee until submission of the Missing Parts of Application.

DO NOT CHARGE THE FILING FEE AT THIS TIME.

Kevin L. Bastian
Reg No.: 34,774
Attorneys for Applicant

Telephone:
(415) 576-0200

Facsimile:
(415) 576-0300

CHLAMYDIA PNEUMONIAE GENOME SEQUENCE

CROSS-REFERENCES TO RELATED APPLICATIONS

The present application is related to 60/128,606, filed April 8, 1999 and
5 60/108,279, filed November 12, 1998, which are incorporated herein by reference.

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

FIELD OF THE INVENTION

This invention relates to nucleic acids and polypeptides from *Chlamydia*
pneumoniae and to their use in the diagnosis, prevention and treatment of diseases
associated with *C. pneumoniae*.

BACKGROUND OF THE INVENTION

Chlamydiaceae is a family of obligate intracellular parasite with a tropism
for epithelial cells lining the mucus membranes. The bacteria have two morphologically
distinct forms, "elementary body" and "reticulate body". The elementary body is the
infectious form, and has a rigid cell wall, primarily of cross-linked outer membrane
20 proteins. The reticulate body is the intracellular, metabolically active form. A unique
developmental cycle between these two forms characterizes *Chlamydia* growth.

C. pneumoniae is a human respiratory pathogen that causes acute
respiratory disease, and approximately 10% of community-acquired pneumonia.
Antibody prevalence studies have shown that virtually everyone is infected with *C.*
25 *pneumoniae* at some time, and that reinfection is common. In addition to respiratory
disease, studies have shown an association of this organism with coronary artery disease.
It has been demonstrated in atherosclerotic lesions of the aorta and coronary arteries by
immunocytochemistry and by polymerase chain reaction (Kuo *et al.* (1993) J Infect Dis
167(4):841-849).

30 Recent reports have further demonstrated the presence of *C. pneumoniae*
in the walls of abdominal aortic aneurysms (Juvonen *et al.* (1997) J Vasc Surg
25(3):499-505). Abdominal aortic aneurysms are frequently associated with
atherosclerosis, and inflammation may be an important factor in aneurysmal dilatation.

C. pneumoniae may play a role in maintaining an inflammation and triggering the development of aortic aneurysms.

Muhlestein *et al.* (1996) JACC 27:1555-61, reported a differential incidence of *Chlamydia* species within the coronary artery wall of patients with atherosclerosis versus those with other forms of cardiovascular disease. The extremely high rate of possible infection in patients with symptomatic atherosclerotic disease compared to the very low rate in patients with normal coronary arteries or coronary artery disease from chronic transplant rejection provides evidence for a direct link between the atherosclerotic process and *Chlamydia* infection. Because a history of chlamydial infection is so prevalent in the population, the issue of causality remains. On a physiologic and pathologic level, abnormal interactions among endothelial cells, platelets, macrophages and lymphocytes may lead to a cascade of events resulting in acute endothelial damage, thrombosis and repair, chronically leading to the development of atheroma in blood vessels.

C. pneumoniae is related to other *Chlamydia* species, but the level of sequence similarity is relatively low. Very little is known about the biology of this organism, although it appears to be an important human pathogen. Allelic diversity and structural relationships between specific genes of Chlamydial species is described in Kaltenboeck *et al.* (1993) J Bacteriol 175(2):487-502; Gaydos *et al.* (1992) Infect Immun 60(12):5319-5323; Everett *et al.* (1997) Int J Syst Bacteriol 47(2):461-473; and Pudjiatmoko *et al.* (1997) Int J Syst Bacteriol 47(2):425-431.

A number of studies have been published describing methods for detection of *C. pneumoniae*, and for distinguishing between Chlamydial species. Such methods include PCR detection (Rasmussen *et al.* (1992) Mol Cell Probes 6(5):389-394; Holland *et al.* (1990) J Infect Dis 162(4):984-987); a simplified polymerase chain reaction-enzyme immunoassay (Wilson *et al.* (1996) J Appl Bacteriol 80(4):431-438); sequence determination and restriction endonuclease cleavage (Herrmann *et al.* (1996) J Clin Microbiol 34(8):1897-1902).

Antigenic and molecular analyses of different *C. pneumoniae* strains is described in Jantos *et al.* (1997) J Clin Microbiol 35(3):620-623. Some genes of *C. pneumoniae* have been isolated and sequenced. These include the Gro E operon (Kikuta *et al.* (1991) Infect Immun 59(12):4665-4669); the major outer membrane protein Perez *et*

al. (1991) Infect Immun 59(6):2195-2199; the DnaK protein homolog (Kornak *et al.* (1991) Infect Immun 59(2):721-725); as well as a number of ribosomal and other genes.

5

SUMMARY OF THE INVENTION

This invention provides the genomic sequence of *Chlamydia pneumoniae*. The sequence information is useful for a variety of diagnostic and analytical methods. The genomic sequence may be embodied in a variety of media, including computer readable forms, or as a nucleic acid comprising a selected fragment of the sequence. Such fragments generally consist of an open reading frame, transcriptional or translational control elements, or fragments derived therefrom. Proteins encoded by the open reading frames are useful for diagnostic purposes, as well as for their enzymatic or structural activity.

10

15

DEFINITIONS

The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group., e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

20

25

30

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

"Amplification" primers are oligonucleotides comprising either natural or analogue nucleotides that can serve as the basis for the amplification of a select nucleic acid sequence. They include, e.g., polymerase chain reaction primers and ligase chain reaction oligonucleotides.

5 "Antibody" refers to an immunoglobulin molecule able to bind to a specific epitope on an antigen. Antibodies can be a polyclonal mixture or monoclonal. Antibodies can be intact immunoglobulins derived from natural sources or from recombinant sources and can be immunoreactive portions of intact immunoglobulins. Antibodies may exist in a variety of forms including, for example, Fv, F_{ab}, and F(ab)₂, as
10 well as in single chains. Single-chain antibodies, in which genes for a heavy chain and a light chain are combined into a single coding sequence, may also be used.

An "antigen" is a molecule that is recognized and bound by an antibody, e.g., peptides, carbohydrates, organic molecules, or more complex molecules such as glycolipids and glycoproteins. The part of the antigen that is the target of antibody
15 binding is an antigenic determinant and a small functional group that corresponds to a single antigenic determinant is called a hapten.

"Biological sample" refers to any sample obtained from a living or dead organism. Examples of biological samples include biological fluids and tissue specimens. Such biological samples can be prepared for analysis of the presence of *C. pneumoniae*
20 nucleic acids, proteins, or antibodies specifically reactive with the proteins.

The term "*C. pneumoniae* gene" shall be intended to mean the open reading frame encoding specific *C. pneumoniae* polypeptides, as well as adjacent 5' and 3' non-coding nucleotide sequences involved in the regulation of expression, up to about 2 kb beyond the coding region, but possibly further in either direction. The gene may be
25 introduced into an appropriate vector for extrachromosomal maintenance or for integration into a host genome.

"Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially
30 identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues

(Batzner *et al.*, *Nucleic Acid Res.* 19:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem.* 260:2605-2608 (1985); Rossolini *et al.*, *Mol. Cell. Probes* 8:91-98 (1994)). Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all
 5 encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the
 10 nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence.

15 As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid.
 20 Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

The following groups each contain amino acids that are conservative substitutions for one another:

- 25 1) Alanine (A), Glycine (G);
- 2) Serine (S), Threonine (T);
- 3) Aspartic acid (D), Glutamic acid (E);
- 4) Asparagine (N), Glutamine (Q);
- 5) Cysteine (C), Methionine (M);
- 30 6) Arginine (R), Lysine (K), Histidine (H);
- 7) Isoleucine (I), Leucine (L), Valine (V); and
- 8) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

see, e.g., Creighton, Proteins (1984)).

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence over a comparison window, as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. This definition also refers to the complement of a test sequence, which has a designated percent sequence or subsequence complementarity when the test sequence has a designated or substantial identity to a reference sequence. For example, a designated amino acid percent identity of 95% refers to sequences or subsequences that have at least about 95% amino acid identity when aligned for maximum correspondence over a comparison window as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Such sequences would then be said to have substantial identity, or to be substantially identical to each other. Preferably, sequences have at least about 70% identity, more preferably 80% identity, more preferably 90-95% identity and above. Preferably, the percent identity exists over a region of the sequence that is at least about 25 amino acids in length, more preferably over a region that is 50-100 amino acids in length.

When percentage of sequence identity is used in reference to proteins or peptides, it is recognized that residue positions that are not identical often differ by conservative amino acid substitutions, where amino acids residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. Where sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Means for making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated according to, e.g., the algorithm of Meyers & Miller, *Computer Applic. Biol. Sci.* 4:11-17 (1988) e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California, USA)..

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated or default program parameters.

A comparison window includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 25 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (*see, e.g., Ausubel et al., supra*).

One example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments to show relationship and percent sequence identity. It also plots a tree or dendrogram showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng & Doolittle, *J. Mol. Evol.* 35:351-360 (1987). The method used is similar to the method described by Higgins & Sharp, *CABIOS* 5:151-153 (1989). The program can align up to 300 sequences, each of a maximum length of 5,000 nucleotides or amino acids. The multiple alignment procedure begins with the pairwise alignment of the two most similar sequences, producing a cluster of two aligned sequences. This cluster is then aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences are aligned by a simple extension of the pairwise alignment of two individual sequences. The

final alignment is achieved by a series of progressive, pairwise alignments. The program is run by designating specific sequences and their amino acid or nucleotide coordinates for regions of sequence comparison and by designating the program parameters. Using PILEUP, a reference sequence is compared to other test sequences to determine the percent sequence identity relationship using the following parameters: default gap weight (3.00), default gap length weight (0.10), and weighted end gaps. PILEUP can be obtained from the GCG sequence analysis software package, e.g, version 7.0 (Devereaux *et al.*, *Nuc. Acids Res.* 12:387-395 (1984).

Another example of algorithm that is suitable for determining percent sequence identity (i.e., substantial similarity or identity) is the BLAST algorithm, which is described in Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues, always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as default parameters a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)).

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (*see, e.g., Karlin & Altschul, Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described below.

Another indication that polynucleotide sequences are substantially identical is if two molecules hybridize to each other under stringent conditions. Stringent conditions are sequence dependent and will be different in different circumstances. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically stringent conditions for a Southern blot protocol involve hybridizing in a buffer comprising 5x SSC, 1% SDS at 65°C or hybridizing in a buffer containing 5x SSC and 1% SDS at 42°C and washing at 65°C with a 0.2x SSC, 0.1% SDS wash.

A "label" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include ³²P, fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, dioxigenin, or haptens and proteins for which antisera or monoclonal antibodies are available.

The term "nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form. The term encompasses nucleic acids containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, which
5 have similar binding properties as the reference nucleic acid, and which are metabolized in a manner similar to the reference nucleotides. Examples of such analogs include, without limitation, phosphorothioates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, peptide-nucleic acids (PNAs).

Unless otherwise indicated, a particular nucleic acid sequence also
10 implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. The term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide, and polynucleotide.

As used herein a "nucleic acid probe or oligonucleotide" is defined as a
15 nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (i.e., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester
20 bond, so long as it does not interfere with hybridization. Thus, for example, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. It will be understood by one of skill in the art that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably
25 directly labeled as with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled such as with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence.

A labeled nucleic acid probe or oligonucleotide is one that is bound, either
30 covalently, through a linker, or through ionic, van der Waals or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe.

"Pharmaceutically acceptable" means a material that is not biologically or otherwise undesirable, i.e., the material can be administered to an individual along with a *Chlamydia* antigen without causing any undesirable biological effects or interacting in a deleterious manner with any of the other components of the pharmaceutical composition.

5 The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an analog or mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers.

10 The phrase "specifically or selectively hybridizing to," refers to hybridization between a probe and a target sequence in which the probe binds substantially only to the target sequence, forming a hybridization complex, when the target is in a heterogeneous mixture of polynucleotides and other compounds. Such hybridization is determinative of the presence of the target sequence. Although the probe
15 may bind other unrelated sequences, at least 90%, preferably 95% or more of the hybridization complexes formed are with the target sequence.

The term "recombinant" when used with reference to a cell, or nucleic acid, or vector, indicates that the cell, or nucleic acid, or vector, has been modified by the introduction of a heterologous nucleic acid or the alteration of a native nucleic acid, or
20 that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

The phrase "specifically immunoreactive with", when referring to a protein
25 or peptide, refers to a binding reaction between the protein and an antibody which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins and other compounds. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein and do not bind in a significant amount to other proteins present in the sample. Specific binding to an antibody under such
30 conditions may require an antibody that is selected for its specificity for a particular protein. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein and are described in detail below.

The phrase "substantially pure" or "isolated" when referring to a *Chlamydia* peptide or protein, means a chemical composition which is free of other subcellular components of the *Chlamydia* organism. Typically, a monomeric protein is substantially pure when at least about 85% or more of a sample exhibits a single polypeptide backbone. Minor variants or chemical modifications may typically share the same polypeptide sequence. Depending on the purification procedure, purities of 85%, and preferably over 95% pure are possible. Protein purity or homogeneity may be indicated by a number of means well known in the art, such as polyacrylamide gel electrophoresis of a protein sample, followed by visualizing a single polypeptide band on a polyacrylamide gel upon silver staining. For certain purposes high resolution will be needed and HPLC or a similar means for purification utilized.

DETAILED DESCRIPTION

The present invention provides the nucleotide sequence of the *C. pneumoniae* genome SEQ ID NO: 1 or a representative fragment thereof, in a form which can be readily used, analyzed, and interpreted by a skilled artisan. As used herein, a "representative fragment" of the nucleotide sequence depicted in SEQ ID NO: 1 refers to any portion which is not presently represented within a publicly available database. Preferred representative fragments of the present invention are open reading frames, expression modulating fragments, uptake modulating fragments, and fragments which can be used to diagnose the presence of *C. pneumoniae* in sample. Using the information provided in the present application, together with routine cloning and sequencing methods, one of ordinary skill in the art will be able to clone and sequence all "representative fragments" of interest including open reading frames (ORFs) encoding a large variety of *C. pneumoniae* proteins. A non-limiting identification of such preferred representative fragments is provided in Tables 2 and 3.

Diagnostic use of *C. pneumoniae* nucleic acids

Hybridization-based assays

Using the nucleic acids disclosed here, one of skill can design nucleic acid hybridization-based assays for the detection of *C. pneumoniae*. Any of a number of well known techniques for the specific detection of target nucleic acids can be used. Exemplary hybridization-based assays include, but are not limited to, traditional "direct

probe" methods such as Southern Blots, dot blots, *in situ* hybridization (*e.g.*, FISH), PCR, and the like. The methods can be used in a wide variety of formats including, but not limited to substrate- (*e.g.* membrane or glass) bound methods or array-based approaches as described below. As noted above, this invention also embraces methods for detecting the presence of *Chlamydia* DNA or RNA in biological samples. These sequences can be used to detect *Chlamydia* in biological samples from patients suspected of being infected. A variety of methods of specific DNA and RNA measurement using nucleic acid hybridization techniques are known to those of skill in the art (*see* Sambrook *et al.*, *supra*).

In situ hybridization assays are well known (*e.g.*, Angerer (1987) *Meth. Enzymol* 152: 649). Generally, *in situ* hybridization comprises the following major steps: (1) fixation of tissue or biological structure to be analyzed; (2) prehybridization treatment of the biological structure to increase accessibility of target DNA, and to reduce nonspecific binding; (3) hybridization of the mixture of nucleic acids to the nucleic acid in the biological structure or tissue; (4) post-hybridization washes to remove nucleic acid fragments not bound in the hybridization and (5) detection of the hybridized nucleic acid fragments. The reagent used in each of these steps and the conditions for use vary depending on the particular application.

In a typical *in situ* hybridization assay, cells are fixed to a solid support, typically a glass slide. If a nucleic acid is to be probed, the cells are typically denatured with heat or alkali. The cells are then contacted with a hybridization solution at a moderate temperature to permit annealing of labeled probes specific to the nucleic acid sequence encoding the protein. The targets (*e.g.*, cells) are then typically washed at a predetermined stringency or at an increasing stringency until an appropriate signal to noise ratio is obtained.

The nucleic acids of this invention are particularly well suited to array-based hybridization formats. Arrays are a multiplicity of different "probe" or "target" nucleic acids (or other compounds) attached to one or more surfaces (*e.g.*, solid, membrane, or gel). In a preferred embodiment, the multiplicity of nucleic acids (or other moieties) is attached to a single contiguous surface or to a multiplicity of surfaces juxtaposed to each other.

In an array format a large number of different hybridization reactions can be run essentially "in parallel." This provides rapid, essentially simultaneous, evaluation

of a number of hybridizations in a single "experiment". Methods of performing hybridization reactions in array based formats are well known to those of skill in the art (*see, e.g.*, Pastinen (1997) *Genome Res.* 7: 606-614; Jackson (1996) *Nature Biotechnology* 14:1685; Chee (1995) *Science* 274: 610; WO 96/17958.

5 Arrays, particularly nucleic acid arrays can be produced according to a wide variety of methods well known to those of skill in the art. For example, in a simple embodiment, "low density" arrays can simply be produced by spotting (*e.g.* by hand using a pipette) different nucleic acids at different locations on a solid support (*e.g.* a glass surface, a membrane, *etc.*).

10 This simple spotting, approach has been automated to produce high density spotted arrays (*see, e.g.*, U.S. Patent No: 5,807,522). This patent describes the use of an automated systems that taps a microcapillary against a surface to deposit a small volume of a biological sample. The process is repeated to generate high density arrays. Arrays can also be produced using oligonucleotide synthesis technology. Thus, for
15 example, U.S. Patent No. 5,143,854 and PCT patent publication Nos. WO 90/15070 and 92/10092 teach the use of light-directed combinatorial synthesis of high density oligonucleotide arrays.

 Many methods for immobilizing nucleic acids on a variety of solid surfaces are known in the art. A wide variety of organic and inorganic polymers, as well
20 as other materials, both natural and synthetic, can be employed as the material for the solid surface. Illustrative solid surfaces include, *e.g.*, nitrocellulose, nylon, glass, quartz, diazotized membranes (paper or nylon), silicones, polyformaldehyde, cellulose, and cellulose acetate. In addition, plastics such as polyethylene, polypropylene, polystyrene, and the like can be used. Other materials which may be employed include paper,
25 ceramics, metals, metalloids, semiconductive materials, cermets or the like. In addition, substances that form gels can be used. Such materials include, *e.g.*, proteins (*e.g.*, gelatins), lipopolysaccharides, silicates, agarose and polyacrylamides. Where the solid surface is porous, various pore sizes may be employed depending upon the nature of the system.

30 In preparing the surface, a plurality of different materials may be employed, particularly as laminates, to obtain various properties. For example, proteins (*e.g.*, bovine serum albumin) or mixtures of macromolecules (*e.g.*, Denhardt's solution) can be employed to avoid non-specific binding, simplify covalent conjugation, enhance

signal detection or the like. If covalent bonding between a compound and the surface is desired, the surface will usually be polyfunctional or be capable of being polyfunctionalized. Functional groups which may be present on the surface and used for linking can include carboxylic acids, aldehydes, amino groups, cyano groups, ethylenic groups, hydroxyl groups, mercapto groups and the like. The manner of linking a wide variety of compounds to various surfaces is well known and is amply illustrated in the literature.

For example, methods for immobilizing nucleic acids by introduction of various functional groups to the molecules is known (*see, e.g.,* Bischoff (1987) *Anal. Biochem.*, 164: 336-344; Kremsky (1987) *Nucl. Acids Res.* 15: 2891-2910). Modified nucleotides can be placed on the target using PCR primers containing the modified nucleotide, or by enzymatic end labeling with modified nucleotides. Use of glass or membrane supports (*e.g.,* nitrocellulose, nylon, polypropylene) for the nucleic acid arrays of the invention is advantageous because of well developed technology employing manual and robotic methods of arraying targets at relatively high element densities. Such membranes are generally available and protocols and equipment for hybridization to membranes is well known.

Target elements of various sizes, ranging from 1 mm diameter down to 1 μm can be used. Smaller target elements containing low amounts of concentrated, fixed probe DNA are used for high complexity comparative hybridizations since the total amount of sample available for binding to each target element will be limited. Thus it is advantageous to have small array target elements that contain a small amount of concentrated probe DNA so that the signal that is obtained is highly localized and bright. Such small array target elements are typically used in arrays with densities greater than $10^4/\text{cm}^2$. Relatively simple approaches capable of quantitative fluorescent imaging of 1 cm^2 areas have been described that permit acquisition of data from a large number of target elements in a single image (*see, e.g.,* Wittrup (1994) *Cytometry* 16:206-213).

If fluorescently labeled nucleic acid samples are used, arrays on solid surface substrates with much lower fluorescence than membranes, such as glass, quartz, or small beads, can achieve much better sensitivity. Substrates such as glass or fused silica are advantageous in that they provide a very low fluorescence substrate, and a highly efficient hybridization environment. Covalent attachment of the target nucleic acids to glass or synthetic fused silica can be accomplished according to a number of

known techniques (described above). Nucleic acids can be conveniently coupled to glass using commercially available reagents. For instance, materials for preparation of silanized glass with a number of functional groups are commercially available or can be prepared using standard techniques (see, *e.g.*, Gait (1984) *Oligonucleotide Synthesis: A Practical Approach*, IRL Press, Wash., D.C.). Quartz cover slips, which have at least 10-fold lower autofluorescence than glass, can also be silanized.

Alternatively, probes can also be immobilized on commercially available coated beads or other surfaces. For instance, biotin end-labeled nucleic acids can be bound to commercially available avidin-coated beads. Streptavidin or anti-digoxigenin antibody can also be attached to silanized glass slides by protein-mediated coupling using *e.g.*, protein A following standard protocols (see, *e.g.*, Smith (1992) *Science* 258: 1122-1126). Biotin or digoxigenin end-labeled nucleic acids can be prepared according to standard techniques. Hybridization to nucleic acids attached to beads is accomplished by suspending them in the hybridization mix, and then depositing them on the glass substrate for analysis after washing. Alternatively, paramagnetic particles, such as ferric oxide particles, with or without avidin coating, can be used.

A variety of other nucleic acid hybridization formats are known to those skilled in the art. For example, common formats include sandwich assays and competition or displacement assays. Hybridization techniques are generally described in Hames and Higgins (1985) *Nucleic Acid Hybridization, A Practical Approach*, IRL Press; Gall and Pardue (1969) *Proc. Natl. Acad. Sci. USA* 63: 378-383; and John *et al.* (1969) *Nature* 223: 582-587.

Sandwich assays are commercially useful hybridization assays for detecting or isolating nucleic acid sequences. Such assays utilize a "capture" nucleic acid covalently immobilized to a solid support and a labeled "signal" nucleic acid in solution. The sample will provide the target nucleic acid. The "capture" nucleic acid and "signal" nucleic acid probe hybridize with the target nucleic acid to form a "sandwich" hybridization complex. To be most effective, the signal nucleic acid should not hybridize with the capture nucleic acid.

Detection of a hybridization complex may require the binding of a signal generating complex to a duplex of target and probe polynucleotides or nucleic acids. Typically, such binding occurs through ligand and anti-ligand interactions as between a ligand-conjugated probe and an anti-ligand conjugated with a signal.

The sensitivity of the hybridization assays may be enhanced through use of a nucleic acid amplification system that multiplies the target nucleic acid being detected. Examples of such systems include the polymerase chain reaction (PCR) system and the ligase chain reaction (LCR) system. Other methods recently described in the art are the nucleic acid sequence based amplification (NASBAO, Cangene, Mississauga, Ontario) and Q Beta Replicase systems.

Nucleic acid hybridization simply involves providing a denatured probe and target nucleic acid under conditions where the probe and its complementary target can form stable hybrid duplexes through complementary base pairing. The nucleic acids that do not form hybrid duplexes are then washed away leaving the hybridized nucleic acids to be detected, typically through detection of an attached detectable label. It is generally recognized that nucleic acids are denatured by increasing the temperature or decreasing the salt concentration of the buffer containing the nucleic acids, or in the addition of chemical agents, or the raising of the pH. Under low stringency conditions (*e.g.*, low temperature and/or high salt and/or high target concentration) hybrid duplexes (*e.g.*, DNA:DNA, RNA:RNA, or RNA:DNA) will form even where the annealed sequences are not perfectly complementary. Thus specificity of hybridization is reduced at lower stringency. Conversely, at higher stringency (*e.g.*, higher temperature or lower salt) successful hybridization requires fewer mismatches.

One of skill in the art will appreciate that hybridization conditions may be selected to provide any degree of stringency. In a preferred embodiment, hybridization is performed at low stringency to ensure hybridization and then subsequent washes are performed at higher stringency to eliminate mismatched hybrid duplexes. Successive washes may be performed at increasingly higher stringency (*e.g.*, down to as low as 0.25 X SSPE-T at 37°C to 70°C) until a desired level of hybridization specificity is obtained. Stringency can also be increased by addition of agents such as formamide. Hybridization specificity may be evaluated by comparison of hybridization to the test probes with hybridization to the various controls that can be present.

In general, there is a tradeoff between hybridization specificity (stringency) and signal intensity. Thus, in a preferred embodiment, the wash is performed at the highest stringency that produces consistent results and that provides a signal intensity greater than approximately 10% of the background intensity. Thus, in a preferred embodiment, the hybridized array may be washed at successively higher

stringency solutions and read between each wash. Analysis of the data sets thus produced will reveal a wash stringency above which the hybridization pattern is not appreciably altered and which provides adequate signal for the particular probes of interest.

Methods of optimizing hybridization conditions are well known to those of skill in the art (*see, e.g.*, Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 24: Hybridization With Nucleic Acid Probes*, Elsevier, N.Y.).

Labeling and detection of nucleic acids.

In a preferred embodiment, the hybridized nucleic acids are detected by detecting one or more labels attached to the sample or probe nucleic acids. The labels may be incorporated by any of a number of means well known to those of skill in the art. Means of attaching labels to nucleic acids include, for example nick translation or end-labeling (*e.g.* with a labeled RNA) by kinasing of the nucleic acid and subsequent attachment (ligation) of a nucleic acid linker joining the sample nucleic acid to a label (*e.g.*, a fluorophore). A wide variety of linkers for the attachment of labels to nucleic acids are also known. In addition, intercalating dyes and fluorescent nucleotides can also be used.

Detectable labels suitable for use in the present invention include any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include biotin for staining with labeled streptavidin conjugate, magnetic beads (*e.g.*, Dynabeads™), fluorescent dyes (*e.g.*, fluorescein, texas red, rhodamine, green fluorescent protein, and the like, *see, e.g.*, Molecular Probes, Eugene, Oregon, USA), radiolabels (*e.g.*, ^3H , ^{125}I , ^{35}S , ^{14}C , or ^{32}P), enzymes (*e.g.*, horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric labels such as colloidal gold (*e.g.*, gold particles in the 40 -80 nm diameter size range scatter green light with high efficiency) or colored glass or plastic (*e.g.*, polystyrene, polypropylene, latex, etc.) beads. Patents teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241.

A fluorescent label is preferred because it provides a very strong signal with low background. It is also optically detectable at high resolution and sensitivity through a quick scanning procedure. The nucleic acid samples can all be labeled with a single label, *e.g.*, a single fluorescent label. Alternatively, in another embodiment, different nucleic acid samples can be simultaneously hybridized where each nucleic acid

sample has a different label. For instance, one target could have a green fluorescent label and a second target could have a red fluorescent label. The scanning step will distinguish sites of binding of the red label from those binding the green fluorescent label. Each nucleic acid sample (target nucleic acid) can be analyzed independently from one another.

5 Suitable chromogens which can be employed include those molecules and compounds which absorb light in a distinctive range of wavelengths so that a color can be observed or, alternatively, which emit light when irradiated with radiation of a particular wave length or wave length range, *e.g.*, fluorescers.

10 Desirably, fluorescers should absorb light above about 300 nm, preferably about 350 nm, and more preferably above about 400 nm, usually emitting at wavelengths greater than about 10 nm higher than the wavelength of the light absorbed. It should be noted that the absorption and emission characteristics of the bound dye can differ from the unbound dye. Therefore, when referring to the various wavelength ranges and characteristics of the dyes, it is intended to indicate the dyes as employed and not the dye
15 which is unconjugated and characterized in an arbitrary solvent.

 Fluorescers are generally preferred because by irradiating a fluorescer with light, one can obtain a plurality of emissions. Thus, a single label can provide for a plurality of measurable events.

20 Detectable signal can also be provided by chemiluminescent and bioluminescent sources. Chemiluminescent sources include a compound which becomes electronically excited by a chemical reaction and can then emit light which serves as the detectable signal or donates energy to a fluorescent acceptor. Alternatively, luciferins can be used in conjunction with luciferase or lucigenins to provide bioluminescence.

25 Spin labels are provided by reporter molecules with an unpaired electron spin which can be detected by electron spin resonance (ESR) spectroscopy. Exemplary spin labels include organic free radicals, transitional metal complexes, particularly vanadium, copper, iron, and manganese, and the like. Exemplary spin labels include nitroxide free radicals.

30 The label may be added to the target (sample) nucleic acid(s) prior to, or after the hybridization. So called "direct labels" are detectable labels that are directly attached to or incorporated into the target (sample) nucleic acid prior to hybridization. In contrast, so called "indirect labels" are joined to the hybrid duplex after hybridization. Often, the indirect label is attached to a binding moiety that has been attached to the

target nucleic acid prior to the hybridization. Thus, for example, the target nucleic acid may be biotinylated before the hybridization. After hybridization, an avidin-conjugated fluorophore will bind the biotin bearing hybrid duplexes providing a label that is easily detected. For a detailed review of methods of labeling nucleic acids and detecting labeled hybridized nucleic acids see *Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 24: Hybridization With Nucleic Acid Probes*, P. Tijssen, ed. Elsevier, N.Y., (1993)).

Fluorescent labels are easily added during an *in vitro* transcription reaction. Thus, for example, fluorescein labeled UTP and CTP can be incorporated into the RNA produced in an *in vitro* transcription.

The labels can be attached directly or through a linker moiety. In general, the site of label or linker-label attachment is not limited to any specific position. For example, a label may be attached to a nucleoside, nucleotide, or analogue thereof at any position that does not interfere with detection or hybridization as desired. For example, certain Label-ON Reagents from Clontech (Palo Alto, CA) provide for labeling interspersed throughout the phosphate backbone of an oligonucleotide and for terminal labeling at the 3' and 5' ends. As shown for example herein, labels can be attached at positions on the ribose ring or the ribose can be modified and even eliminated as desired. The base moieties of useful labeling reagents can include those that are naturally occurring or modified in a manner that does not interfere with the purpose to which they are put. Modified bases include but are not limited to 7-deaza A and G, 7-deaza-8-aza A and G, and other heterocyclic moieties.

It will be recognized that fluorescent labels are not to be limited to single species organic molecules, but include inorganic molecules, multi-molecular mixtures of organic and/or inorganic molecules, crystals, heteropolymers, and the like. Thus, for example, CdSe-CdS core-shell nanocrystals enclosed in a silica shell can be easily derivatized for coupling to a biological molecule (Bruchez *et al.* (1998) *Science*, 281: 2013-2016). Similarly, highly fluorescent quantum dots (zinc sulfide-capped cadmium selenide) have been covalently coupled to biomolecules for use in ultrasensitive biological detection (Warren and Nie (1998) *Science*, 281: 2016-2018).

Amplification-based assays.

In another embodiment, amplification-based assays can be used to detect nucleic acids. In such amplification-based assays, the nucleic acid sequences act as a

template in an amplification reaction (e.g. Polymerase Chain Reaction (PCR). Detailed protocols for quantitative PCR are provided in Innis *et al.* (1990) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc. N.Y.).

Other suitable amplification methods include, but are not limited to ligase chain reaction (LCR) (see Wu and Wallace (1989) *Genomics* 4: 560, Landegren *et al.* (1988) *Science* 241: 1077, and Barringer *et al.* (1990) *Gene* 89: 117, transcription amplification (Kwoh *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86: 1173), and self-sustained sequence replication (Guatelli *et al.* (1990) *Proc. Nat. Acad. Sci. USA* 87: 1874).

Detection of *C. pneumoniae* gene expression

The nucleic acids of the invention can also be used to *C. pneumoniae* detect gene transcripts. Methods of detecting and/or quantifying gene transcripts using nucleic acid hybridization techniques are known to those of skill in the art (see Sambrook *et al. supra*). For example, a Northern transfer may be used for the detection of the desired mRNA directly. In brief, the mRNA is isolated from a given cell sample using, for example, an acid guanidinium-phenol-chloroform extraction method. The mRNA is then electrophoresed to separate the mRNA species and the mRNA is transferred from the gel to a nitrocellulose membrane. As with the Southern blots, labeled probes are used to identify and/or quantify the target mRNA.

In another preferred embodiment, the gene transcript can be measured using amplification (e.g. PCR) based methods as described above for directly assessing copy number of the target sequences.

Expression of *C. pneumoniae* proteins

The nucleic acids disclosed here can be used for recombinant expression of the proteins. In these methods, the nucleic acids encoding the proteins of interest are introduced into suitable host cells, followed by induction of the cells to produce large amounts of the protein. The invention relies on routine techniques in the field of recombinant genetics, well known to those of ordinary skill in the art. A basic text disclosing the general methods of use in this invention is Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (2nd ed. 1989).

Standard transfection methods are used to produce prokaryotic, mammalian, yeast or insect cell lines which express large quantities of the desired

polypeptide, which is then purified using standard techniques (*see, e.g., Colley et al., J. Biol. Chem.* 264:17619-17622, 1989; *Guide to Protein Purification, supra*).

The nucleotide sequences used to transfect the host cells can be modified to yield *Chlamydia* polypeptides with a variety of desired properties. For example, the polypeptides can vary from the naturally-occurring sequence at the primary structure level by amino acid, insertions, substitutions, deletions, and the like. These modifications can be used in a number of combinations to produce the final modified protein chain.

The amino acid sequence variants can be prepared with various objectives in mind, including facilitating purification and preparation of the recombinant polypeptide. The modified polypeptides are also useful for modifying plasma half life, improving therapeutic efficacy, and lessening the severity or occurrence of side effects during therapeutic use. The amino acid sequence variants are usually predetermined variants not found in nature but exhibit the same immunogenic activity as naturally occurring protein. In general, modifications of the sequences encoding the polypeptides may be readily accomplished by a variety of well-known techniques, such as site-directed mutagenesis (*see Gillman & Smith, Gene* 8:81-97 (1979); Roberts *et al., Nature* 328:731-734 (1987)). One of ordinary skill will appreciate that the effect of many mutations is difficult to predict. Thus, most modifications are evaluated by routine screening in a suitable assay for the desired characteristic. For instance, the effect of various modifications on the ability of the polypeptide to elicit a protective immune response can be easily determined using *in vitro* assays. For instance, the polypeptides can be tested for their ability to induce lymphoproliferation, T cell cytotoxicity, or cytokine production using standard techniques.

The particular procedure used to introduce the genetic material into the host cell for expression of the polypeptide is not particularly critical. Any of the well known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection, plasmid vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (*see Sambrook et al., supra*). It is only necessary that the particular procedure utilized be capable of successfully introducing at least one gene into the host cell which is capable of expressing the gene.

Any of a number of well known cells and cell lines can be used to express the polypeptides of the invention. For instance, prokaryotic cells such as *E. coli* can be used. Eukaryotic cells include, yeast, Chinese hamster ovary (CHO) cells, COS cells, and insect cells.

5 The particular vector used to transport the genetic information into the cell is also not particularly critical. Any of the conventional vectors used for expression of recombinant proteins in prokaryotic and eukaryotic cells may be used. Expression vectors for mammalian cells typically contain regulatory elements from eukaryotic viruses.

10 The expression vector typically contains a transcription unit or expression cassette that contains all the elements required for the expression of the polypeptide DNA in the host cells. A typical expression cassette contains a promoter operably linked to the DNA sequence encoding a polypeptide and signals required for efficient polyadenylation of the transcript. The term "operably linked" as used herein refers to linkage of a
15 promoter upstream from a DNA sequence such that the promoter mediates transcription of the DNA sequence. The promoter is preferably positioned about the same distance from the heterologous transcription start site as it is from the transcription start site in its natural setting. As is known in the art, however, some variation in this distance can be accommodated without loss of promoter function.

20 Following the growth of the recombinant cells and expression of the polypeptide, the culture medium is harvested for purification of the secreted protein. The media are typically clarified by centrifugation or filtration to remove cells and cell debris and the proteins are concentrated by adsorption to any suitable resin or by use of ammonium sulfate fractionation, polyethylene glycol precipitation, or by ultrafiltration.
25 Other routine means known in the art may be equally suitable. Further purification of the polypeptide can be accomplished by standard techniques, for example, affinity chromatography, ion exchange chromatography, sizing chromatography, His₆ tagging and Ni-agarose chromatography (as described in Dobeli *et al.*, *Mol. and Biochem. Parasit.* 41:259-268 (1990)), or other protein purification techniques to obtain homogeneity. The
30 purified proteins are then used to produce pharmaceutical compositions, as described below.

 An alternative method of preparing recombinant polypeptides useful as vaccines involves the use of recombinant viruses (e.g., vaccinia). Vaccinia virus is grown

in suitable cultured mammalian cells such as the HeLa S3 spinner cells, as described by Mackett *et al.*, in *DNA cloning Vol. II: A practical approach*, pp. 191-211 (Glover, ed.).

Antibody Production

The proteins of the present invention can be used to produce antibodies specifically reactive with *C pneumoniae* antigens. If isolated proteins are used, they may be recombinantly produced or isolated from *Chlamydia* cultures. Synthetic peptides made using the protein sequences may also be used.

Methods of production of polyclonal antibodies are known to those of skill in the art. In brief, an immunogen, preferably a purified protein, is mixed with an adjuvant and animals are immunized. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera is prepared. Further fractionation of the antisera to enrich for antibodies reactive to *Chlamydia* proteins can be done if desired (*see* Harlow & Lane, *Antibodies: A Laboratory Manual* (1988)).

Polyclonal antisera are used to identify and characterize *Chlamydia* in the tissues of patients using, for instance, *in situ* techniques and immunoperoxidase test procedures described in Anderson *et al. JAVMA* 198:241 (1991) and Barr *et al. Vet. Pathol.* 28:110-116 (1991).

Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (*see* Kohler & Milstein, *Eur. J. Immunol.* 6:511-519 (1976)). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host.

Monoclonal antibodies produced in such a manner are used, for instance, in ELISA diagnostic tests, immunoperoxidase tests, immunohistochemical tests, for the *in vitro* evaluation of spirochete invasion, to select candidate antigens for vaccine development, protein isolation, and for screening genomic and cDNA libraries to select appropriate gene sequences.

Immunodiagnostic detection of *C. pneumoniae* infections

The present invention also provides methods for detecting the presence or absence of *C. pneumoniae*, or antibodies reactive with it, in a biological sample. For instance, antibodies specifically reactive with *Chlamydia* can be detected using either
5 *Chlamydia* proteins or the isolates described here. The proteins and isolates can also be used to raise specific antibodies (either monoclonal or polyclonal) to detect the antigen in a sample. In addition, the nucleic acids disclosed and claimed here can be used to detect *Chlamydia*-specific sequences using standard hybridization techniques.

For a review of immunological and immunoassay procedures in general,
10 see *Basic and Clinical Immunology* (Stites & Terr ed., 7th ed. 1991)). The immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in *Enzyme Immunoassay* (Maggio, ed., 1980); Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology* (1985)). For instance, the proteins and antibodies disclosed here are conveniently used in ELISA, immunoblot analysis and
15 agglutination assays.

In brief, immunoassays to measure anti-*Chlamydia* antibodies or antigens can be either competitive or noncompetitive binding assays. In competitive binding assays, the sample analyte (e.g., anti-*Chlamydia* antibodies) competes with a labeled analyte (e.g., anti-*Chlamydia* monoclonal antibody) for specific binding sites on a capture
20 agent (e.g., isolated *Chlamydia* protein) bound to a solid surface. The concentration of labeled analyte bound to the capture agent is inversely proportional to the amount of free analyte present in the sample.

Noncompetitive assays are typically sandwich assays, in which the sample analyte is bound between two analyte-specific binding reagents. One of the binding
25 agents is used as a capture agent and is bound to a solid surface. The second binding agent is labelled and is used to measure or detect the resultant complex by visual or instrument means.

A number of combinations of capture agent and labelled binding agent can be used. For instance, an isolated *Chlamydia* protein or culture can be used as the
30 capture agent and labelled anti-human antibodies specific for the constant region of human antibodies can be used as the labelled binding agent. Goat, sheep and other non-human antibodies specific for human immunoglobulin constant regions (e.g., γ or μ) are

well known in the art. Alternatively, the anti-human antibodies can be the capture agent and the antigen can be labelled.

Various components of the assay, including the antigen, anti-*Chlamydia* antibody, or anti-human antibody, may be bound to a solid surface. Many methods for immobilizing biomolecules to a variety of solid surfaces are known in the art. For instance, the solid surface may be a membrane (e.g., nitrocellulose), a microtiter dish (e.g., PVC or polystyrene) or a bead. The desired component may be covalently bound or noncovalently attached through nonspecific bonding.

Alternatively, the immunoassay may be carried out in liquid phase and a variety of separation methods may be employed to separate the bound labeled component from the unbound labelled components. These methods are known to those of skill in the art and include immunoprecipitation, column chromatography, adsorption, addition of magnetizable particles coated with a binding agent and other similar procedures.

An immunoassay may also be carried out in liquid phase without a separation procedure. Various homogeneous immunoassay methods are now being applied to immunoassays for protein analytes. In these methods, the binding of the binding agent to the analyte causes a change in the signal emitted by the label, so that binding may be measured without separating the bound from the unbound labelled component.

Western blot (immunoblot) analysis can also be used to detect the presence of antibodies to *Chlamydia* in the sample. This technique is a reliable method for confirming the presence of antibodies against a particular protein in the sample. The technique generally comprises separating proteins by gel electrophoresis on the basis of molecular weight, transferring the separated proteins to a suitable solid support, (such as a nitrocellulose filter, a nylon filter, or derivatized nylon filter), and incubating the sample with the separated proteins. This causes specific target antibodies present in the sample to bind their respective proteins. Target antibodies are then detected using labeled anti-human antibodies.

The immunoassay formats described above employ labelled assay components. The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. A wide variety of labels may be used. The component may be labelled by any one of several methods. Traditionally a radioactive label incorporating ^3H , ^{125}I , ^{35}S , ^{14}C , or ^{32}P was used. Non-radioactive labels

include ligands which bind to labelled antibodies, fluorophores, chemiluminescent agents, enzymes, and antibodies which can serve as specific binding pair members for a labelled ligand. The choice of label depends on sensitivity required, ease of conjugation with the compound, stability requirements, and available instrumentation.

Enzymes of interest as labels will primarily be hydrolases, particularly phosphatases, esterases and glycosidases, or oxidoreductases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescent compounds include luciferin, and 2,3-dihydrophthalazinediones, e.g., luminol. For a review of various labelling or signal producing systems which may be used, see U.S. Patent No. 4,391,904, which is incorporated herein by reference.

Non-radioactive labels are often attached by indirect means. Generally, a ligand molecule (e.g., biotin) is covalently bound to the molecule. The ligand then binds to an anti-ligand (e.g., streptavidin) molecule which is either inherently detectable or covalently bound to a signal system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. A number of ligands and anti-ligands can be used. Where a ligand has a natural anti-ligand, for example, biotin, thyroxine, and cortisol, it can be used in conjunction with the labelled, naturally occurring anti-ligands. Alternatively, any haptenic or antigenic compound can be used in combination with an antibody.

Some assay formats do not require the use of labelled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples comprising the target antibodies. In this format, none of the components need be labelled and the presence of the target antibody is detected by simple visual inspection.

Pharmaceutical Compositions

The peptides or antibodies (typically monoclonal antibodies) of the present invention and pharmaceutical compositions thereof are useful for administration to mammals, particularly humans, to treat and/or prevent *Chlamydia* infections. Suitable formulations are found in *Remington's Pharmaceutical Sciences*, Mack Publishing Company, Philadelphia, PA, 17th ed. (1985).

The immunogenic peptides or antibodies of the invention are administered prophylactically or to an individual already suffering from the disease. The peptide compositions are administered to a patient in an amount sufficient to elicit an effective immune response to *Chlamydia*. An effective immune response is one that inhibits infection. An amount adequate to accomplish this is defined as "therapeutically effective dose" or "immunogenically effective dose." Amounts effective for this use will depend on, e.g., the peptide composition, the manner of administration, the stage and severity of the disease being treated, the weight and general state of health of the patient, and the judgment of the prescribing physician, but generally range for the initial immunization (that is for therapeutic or prophylactic administration) from about 0.1 mg to about 1.0 mg per 70 kilogram patient, more commonly from about 0.5 mg to about 0.75 mg per 70 kg of body weight. Boosting dosages are typically from about 0.1 mg to about 0.5 mg of peptide using a boosting regimen over weeks to months depending upon the patient's response and condition. A suitable protocol would include injection at time 0, 4, 2, 6, 10 and 14 weeks, followed by further booster injections at 24 and 28 weeks.

For therapeutic use, administration should begin at the first sign of infection. This is followed by boosting doses until at least symptoms are substantially abated and for a period thereafter. In some circumstances, loading doses followed by boosting doses may be required. The resulting immune response helps to cure or at least partially arrest symptoms and/or complications. Vaccine compositions containing the peptides are administered prophylactically to a patient susceptible to or otherwise at risk of the infection.

The pharmaceutical compositions (containing either peptides or antibodies) are intended for parenteral or oral administration. Preferably, the pharmaceutical compositions are administered parenterally, e.g., subcutaneously, intradermally, or intramuscularly. Thus, the invention provides compositions for parenteral administration which comprise a solution of the immunogenic polypeptides dissolved or suspended in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be used, e.g., water, buffered water, 0.4% saline, 0.3% glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration. The compositions may contain

pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as buffering agents, tonicity adjusting agents, wetting agents and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, etc.

5 The compositions may also comprise carriers to enhance the immune response. Useful carriers are well known in the art, and include, e.g., KLH, thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly(lysine:glutamic acid), influenza, hepatitis B virus core protein, hepatitis B virus recombinant vaccine and the like.

10 For solid compositions, conventional nontoxic solid carriers may be used which include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic composition is formed by incorporating any of the normally employed excipients, such as
15 those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more peptides of the invention, and more preferably at a concentration of 25%-75%.

 As noted above, the peptide compositions are intended to induce an immune response to *Chlamydia*. Thus, compositions and methods of administration suitable for maximizing the immune response are preferred. For instance, peptides may
20 be introduced into a host, including humans, linked to a carrier or as a homopolymer or heteropolymer of active peptide units from various *Chlamydia* proteins disclosed here. Alternatively, a "cocktail" of polypeptides can be used. A mixture of more than one polypeptide has the advantage of increased immunological reaction and, where different peptides are used to make up the polymer, the additional ability to induce antibodies to a
25 number of epitopes.

 The compositions also include an adjuvant. As used here, number of adjuvants are well known to one skilled in the art. Suitable adjuvants include incomplete Freund's adjuvant, alum, aluminum phosphate, aluminum hydroxide,
N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP),
30 N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (CGP 19835A, referred to as MTP-PE), and RIBI, which contains three components extracted from bacteria, monophosphoryl

lipid A, trehalose dimycolate and cell wall skeleton (MPL+TDM+CWS) in a 2% squalene/Tween 80 emulsion. The effectiveness of an adjuvant may be determined by measuring the amount of antibodies directed against the immunogenic peptide.

The concentration of immunogenic peptides of the invention in the pharmaceutical formulations can vary widely, i.e. from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, etc., in accordance with the particular mode of administration selected.

The peptides of the invention can also be expressed by attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into a host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover et al. (*Nature* 351:456-460 (1991)). A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, e.g., *Salmonella typhi* vectors and the like, will be apparent to those skilled in the art from the description herein.

The DNA encoding one or more of the peptides of the invention can also be administered to the patient. This approach is described, for instance, in Wolff *et. al.*, *Science* 247: 1465-1468 (1990) as well as U.S. Patent Nos. 5,580,859 and 5,589,466.

In order to enhance serum half-life, the peptides may also be encapsulated, introduced into the lumen of liposomes, prepared as a colloid, or other conventional techniques may be employed which provide an extended serum half-life of the peptides. A variety of methods are available for preparing liposomes, as described in, e.g., Szoka et al., *Ann. Rev. Biophys. Bioeng.* 9:467 (1980), U.S. Pat. Nos. 4, 235,871, 4,501,728 and 4,837,028.

EXAMPLES

The following examples are offered to illustrate, but not to limit the claimed invention.

Example 1:

This example describes comparison of the *C. pneumoniae* genome disclosed here and the, previously sequenced, *C. trachomatis* genome (Stephens, *et al. Science* **282**:754-759 (1998)).

The apparent low level of DNA homology between *C. trachomatis* and *C. pneumoniae* (Campbell, *et al., J. Clin. Microbiol.* **25**:1911-1916 (1987)) yet analogous cell structures and developmental cycles, predicts that comparative analysis of the two genomes will significantly enhance the understanding of both pathogens. Identification of genes that are present in one species but not the other are of particular importance for the mutually exclusive biological, virulence and pathogenesis capabilities of each.

Identification of genes shared between the two species strongly supports the requirement for these capabilities in a biological system that has, over its long-term association with mammalian host cells, evolved to reduce the metabolic capacities while optimizing survival, growth and transmission of these unique pathogens.

The previously sequenced *C. trachomatis* genome contains 1,042,519 nucleotides and 875 likely protein-coding genes. Similarity searching permitted the inferred functional assignment of sequences 636 (60%) genes disclosed here and 251 (23%) are similar to hypothetical genes for other bacterial organisms including those for *C. trachomatis*. The remaining 186 (17%) genes are not homologous to sequences deposited in GenBank.. Seventy *C. trachomatis* genes are not represented in the *C. pneumoniae* genome. These are contained within blocks consisting of 2-17 genes and 19 single genes. Of the 70 *C. trachomatis* genes without homologs in *C. pneumoniae*, 60 are classified as encoding hypothetical proteins. The remaining genes not represented in *C. pneumoniae* consist of the tryptophan operon (*trpA,B,R*), *trpC*, two predicted thiol protease genes, and 4 genes assigned to the phospholipase-D superfamily.

It is evident that there is a high level of functional conservation between *C. pneumoniae* and *C. trachomatis* as orthologs to *C. trachomatis* genes were identified for 859 (80%) of the predicted coding sequences for *C. pneumoniae*. The level of similarity for individual encoded proteins spans a wide spectrum (22-95% amino acid identity) with an average of 62% amino acid identity between orthologs from the two species. The percent amino acid identity between orthologous chlamydial proteins is similar among functional groups with the highest for proteins associated with translation and the lowest for proteins whose function in chlamydiae is uncharacterized and not related to proteins encoded by other organisms. The gene order of the homologous set of genes in *C.*

pneumoniae shows reorganization relative to the genome of *C. trachomatis*; however, there is a high level of synteny for the gene organization of the two genomes. We identified thirty-nine blocks of 2 or more genes whose gene organization is colinear with homologs to *C. trachomatis*, although some of these are inverted. The distribution of genome reorganization is not evenly distributed on the chromosome as the region between *C. pneumoniae* coding sequences 0130-0300 contains substantially more reorganization than other areas of the genome. This region coincides with the predicted chromosome replication terminus.

We identified orthologs of enzymes characterized in other bacteria that account for the essential requirements for DNA replication, repair, transcription and translation including two predicted DNA helicases of the Swi2/Snf2 family found in *C. trachomatis*. Similar to *C. trachomatis*, alternative sigma subunits for RNA polymerase, σ^{28} and σ^{54} , were identified in addition to anti- σ regulatory system factors RsbV, a RsbW-like single-domain histidine kinase, and a RsbU-like protein phosphatase. These findings suggest that the fundamental mechanisms of transcriptional regulation are conserved among *Chlamydia*. The *C. trachomatis* proteins containing SET and SWIB domains, and a SWIB domain fused to the C-terminus of the chlamydial topoisomerase I, not identified outside eukaryotes, are found in *C. pneumoniae* supporting their possible role in the chromatin condensation-decondensation characteristic of the biologically unique chlamydial developmental cycle.

The central metabolic pathways inferred from the *C. pneumoniae* genome sequence are the same as those identified for *C. trachomatis*. *C. pneumoniae* has a glycolytic pathway and a linked tricarboxylic acid cycle, although likely functional, is incomplete as genes for citrate synthase, aconitase, and isocitrate dehydrogenase were not identified. *C. pneumoniae* has a complete glycogen synthesis and degradation system supporting a role for glycogen synthesis and utilization of glucose-derivatives in chlamydial metabolism. Genes encoding essential functions in aerobic respiration are present and electron flux may be supported by pyruvate, succinate, glycerol-3-phosphate, and NADH dehydrogenases, NADH-ubiquinone oxidoreductase and cytochrome oxidase. *C. pneumoniae* also contains the V (vacuolar)-type ATPase operon and the two ATP translocases found in *C. trachomatis*.

The type-III secretion virulence system required for invasion by several pathogenic bacteria and found in the *C. trachomatis* genome in three chromosomal

locations is also present in the *C. pneumoniae* genome. Each of the components is conserved and their relative genomic contexts are conserved. Genes such as a predicted serine/threonine protein kinase and other genes physically linked to genes encoding structural components of the type-III secretion apparatus, but without identified
5 homologs, are also highly similar between the two species suggesting the functional roles in modifying cellular biology are fundamentally conserved.

Chlamydia-encoded proteins that are not found in chlamydial organisms but localized to the intracellular chlamydial inclusion membrane are likely essential for the unique intracellular biology and perhaps differences in inclusion morphology
10 observed between species of *Chlamydia*. Several such proteins, termed IncA,B&C, have been characterized for a *C. psittaci* strain (Rockey, *et al. Mol. Microbiol.* 15:617-626 (1995); Rockey *et al. Infect. Immun.* 62:106-112 (1994)). *C. pneumoniae* and *C. trachomatis* encode orthologs to *C. psittaci* IncB and IncC and *C. trachomatis* also contains an ortholog to IncA. *C. pneumoniae* contains two genes that encode proteins
15 with similarity to IncA (CPn0186 and CPn0585), although the level of homology is low suggesting analogous but possibly altered functions.

The tryptophan biosynthesis operon (*trpA*, *trpB*, *trpR*) and *trpC* identified in *C. trachomatis* is conspicuously missing in the *C. pneumoniae* genome. This represents the entire repertoire of genes associated with tryptophan biosynthesis identified
20 in *C. trachomatis*. Seventeen genes adjacent to the *C. trachomatis* tryptophan operon also were not found in the *C. pneumoniae* genome. This region is the single largest loss of a contiguous genomic segment and includes 4 HKD superfamily encoding genes that encompass a family of proteins related to endonuclease and phospholipase D. These findings may be important for the ability of *Chlamydia* to persist in their hosts and cause
25 disease by eliciting potent, focal and persistent inflammatory responses thought to be essential for pathogenesis.

The *C. pneumoniae* genome contains 187,711 additional nucleotides compared to the *C. trachomatis* genome, and the 214 coding sequences not found in *C. trachomatis* account for most of the increased genome size. Eighty-eight of these genes
30 are found in blocks of >10 genes (11-30 genes/block), 41 are single genes, and the remainder are partnered with at least one other gene. Based upon the observation that ~70% of all the *C. pneumoniae* genes have an identifiable homolog in GenBank, exclusive of *C. trachomatis*, it would be expected that over 150 of the 214 genes should

have a homolog in GenBank, many associated with a function. However, only 28 coding sequences have similarity to genes from other organisms. Thus the majority of the genes that are mutually exclusive of *C. trachomatis* (186 of 214), and the 60 of 70 *C.*

trachomatis genes that lacked an identifiable homolog in *C. pneumoniae*, do not have detectable homologs to genes from other organisms. We predict that most of the unique genes are essential for specific attributes that define the differential biology, tropism and pathogenesis of *C. trachomatis* and *C. pneumoniae*. Moreover, this suggests that *C. pneumoniae* has more unique biological (i.e., virulence) capacity than *C. trachomatis*. The ability of *C. pneumoniae* to be more invasive and survive in a broader range of host cell types than *C. trachomatis* is consistent with this hypothesis. Not all of the differences in biological capacity may be associated with mutually exclusive genes. One explanation for the significantly lower level of homology between protein sequences assigned as having *C. pneumoniae* and *C. trachomatis* orthologs but no identifiable orthologs in other organisms is that this set of proteins is not only associated with biological requirements specific for *Chlamydia* but this polymorphism may account for differential biology between the two species. The determination of the genome sequence from a representative of the *C. psittaci* group will precisely delineate those genes that are mutually exclusive and specific for each species.

The major functionally identifiable addition to the *C. pneumoniae* genome is a large expansion of genes encoding a new family of chlamydial polymorphic membrane proteins (Pmp), alone representing 22% of the increased coding capacity. While the *C. trachomatis* genome has 9 *pmp* genes, remarkably the *C. pneumoniae* genome contains 21 *pmp* genes. Most of these genes appear to be amplified in two regions of the genome with three stand-alone genes. Interestingly one of the stand-alone genes is most closely related to the *C. trachomatis pmpD* which is the only stand-alone *pmp* gene in the *C. trachomatis* genome and it is located with the same relative genomic context, suggesting an essential and conserved function for this paralog. Six Pmp-coding genes are presumably not functional as five contain predicted coding frame-shifts and one is truncated. The amplification of this gene family and the confidently predicted frame-shifts suggest a specific molecular mechanism to promote functional or antigenic diversity. The biological role of this protein family remains enigmatic, although at least one of the proteins in *C. psittaci* related to this family is exposed on the chlamydial surface.

While a function could not be assigned for most of the unique *C. pneumoniae* genes, several have significant similarity to genes from other organisms. Functional assignments could be made for genes encoding GMP synthetase, IMP dehydrogenase, UMP synthase, uridine kinase, biotin synthase pathway proteins, methylthioadenosine nucleosidase, a DNA glycosylase and aromatic amino acid hydroxylase. Thus a complete pathway was identified for biotin biosynthesis. The additional purine and pyrimidine salvage pathway genes presumably reflect metabolic limitations in one of the cell types that *C. pneumoniae* infects or differences in the ability of *C. pneumoniae* to transport precursor nucleosides or nucleotides.

The addition of aromatic amino acid hydroxylase in *C. pneumoniae* is intriguing especially in light of the loss of tryptophan biosynthetic genes and the inability to synthesize other amino acids including phenylalanine. Aromatic amino acid hydroxylases include three distinct enzymes that function to receptively oxidize phenylalanine to tyrosine, tyrosine to Dopa, and tryptophan to 5-hydroxytryptophan and serotonin. Although the chlamydial protein is similar to proteins of this family and incrementally more closely related to tryptophan hydroxylase, its specific function could not be confidently predicted. We hypothesize that it may be involved in *C. pneumoniae* virulence. Tryptophan hydroxylase has not been previously identified in bacteria and the origin of the chlamydial gene appears to be from eukaryotes. The functional role of an aromatic amino acid hydroxylase for *C. pneumoniae* is linked to the unique intracellular biology of this organism and may represent a key contribution to *C. pneumoniae* persistence and pathogenesis.

It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

Table 1 provides functional assignments of *C. pneumoniae* nonprotein-encoding genomic sequences. Table 2 provides functional assignments of protein coding sequences. Table 3 provides the amino acid sequences of the proteins corresponding to the coding sequences.

TABLE 1

type	SEQ ID NO:1 start position	SEQ ID NO:1 end position	Gene
Ori	841664	841396	(R) Putative Origin of Replication
tmRNA	138493	138074	(R) tmRNA
pRNA	607342	607649	Ribonuclease P RNA
rRNA	1000564	1002115	16S rRNA
rRNA	1002415	1005278	23S rRNA
rRNA	1005393	1005509	5S rRNA
tRNA	269070	269142	Ala tRNA_1
tRNA	164318	164389	Asn tRNA
tRNA	296224	296151	(R) Asp tRNA
tRNA	836191	836119	(R) Ala tRNA_2
tRNA	1030533	1030603	Cys tRNA
tRNA	784896	784822	(R) Glu tRNA
tRNA	781680	781610	(R) Gly tRNA_1
tRNA	961536	961607	Gly tRNA_2
tRNA	999949	1000023	His tRNA
tRNA	268992	269065	Ile tRNA
tRNA	672236	672318	Leu tRNA_1
tRNA	680178	680257	Leu tRNA_2
tRNA	715889	715971	Leu tRNA_3
tRNA	739403	739486	Leu tRNA_4
tRNA	1175863	1175944	Leu tRNA_5
tRNA	784994	784922	(R) Lys tRNA
tRNA	843926	843999	Pro tRNA_2
tRNA	409922	409848	(R) Pro tRNA_1
tRNA	631373	631445	Phe tRNA
tRNA	677337	677264	(R) Arg tRNA_2
tRNA	807413	807341	(R) Arg tRNA_3
tRNA	877473	877400	(R) Arg tRNA_4
tRNA	462141	462214	Arg tRNA_1
tRNA	1085605	1085676	Gln tRNA
tRNA	786780	786708	(R) Thr tRNA_3
tRNA	89728	89657	(R) Thr tRNA_1
tRNA	293477	293405	(R) Thr tRNA_2
tRNA	87522	87450	(R) Met tRNA_1
tRNA	199301	199229	(R) Met tRNA_2
tRNA	199390	199317	(R) Met tRNA_3
tRNA	626904	626987	Ser tRNA_1
tRNA	708359	708440	Ser tRNA_2
tRNA	1142034	1142117	Ser tRNA_3
tRNA	1230028	1229945	(R) Ser tRNA_4
tRNA	91070	90999	(R) Trp tRNA
tRNA	293399	293317	(R) Tyr tRNA
tRNA	296147	296075	(R) Val tRNA_1
tRNA	1137389	1137462	Val tRNA_2

TABLE 2

Gene #	From	To	Strand	Gene Function (C. trachomatis ortholog in parentheses)
CPn0001	282	4	R	CT001 hypothetical protein
CPn0002	573	875	F	gatC-Glu-tRNA Gln Amidotransferase (C subunit)-(CT002)
CPn0003	895	2370	F	gatA-Glu tRNA Gln Amidotransferase-(CT003)
CPn0004	2170	1111	F	gatB-(Pef112) Gln tRNA Gln Amidotransferase (B Subunit)-(CT004)
CPn0005	4127	6892	F	pmp_1-Polymorphic Outer Membrane Protein G Family
CPn0006	7293	7141	R	
CPn0007	7605	10496	F	
CPn0008	10975	11685	F	
CPn0009	11815	13119	F	
CPn0010	13435	14325	F	
CPn0010	14379	15746	F	frame-shift with 0010
CPn0011	15892	16614	F	
CPn0012	16644	18212	F	
CPn0013	18584	21106	F	pmp_2-Polymorphic Outer Membrane Protein G Family
CPn0014	21392	21922	F	pmp_3-Polymorphic Outer Membrane Protein G Family
CPn0015	21835	24174	F	pmp_3-PMP_3 (frame-shift with 0014)
CPn0016	24416	26188	F	pmp_4-Polymorphic Outer Membrane Protein G Family
CPn0017	26094	27170	F	pmp_4-PMP_4 (frame-shift with 0016)
CPn0018	27522	29003	F	pmp_5-Polymorphic Outer Membrane Protein G Family
CPn0019	29007	30356	F	pmp_5-PMP_5 (frame-shift with 0018)
CPn0020	32687	30603	R	Predicted OMP [leader (14) peptide: outer membrane]-(CT351)
CPn0021	34410	32707	R	Predicted OMP [leader (19) peptide]-(CT350)
CPn0022	34982	34395	F	maf-(CT349)
CPn0023	36603	35014	R	yjjK/alr-ABC Transporter Protein ATPase-(CT348)
CPn0024	37596	36661	F	xerC-Integrase/recombinase-(CT347)
CPn0025	38604	37684	R	elaC/atsA-Sulphohydrolase/Glycosulfatase-(CT346)
CPn0026	39625	38762	R	CT345 hypothetical protein-(CT345)
CPn0027	42234	39778	R	lon-Lon ATP-dependent Protease-(CT344)
CPn0028	43325	42543	R	
CPn0029	43755	43390	R	
CPn0030	43891	44529	F	gcp_1-O-Sialoglycoprotein Endopeptidase_1-(CT343)
CPn0031	44711	44884	F	rs21-S21 Ribosomal Protein-(CT342)
CPn0032	44923	46098	F	dnaJ-Heat Shock Protein J-(CT341)
CPn0033	46138	48171	F	pdhA&B/odhA&odhB-(pyruvate) Oxoisovalerate Dehydrogenase Alpha & Beta Fusion-(CT340)
CPn0034	49457	48210	R	
CPn0035	51029	49569	R	CT339 hypothetical protein
CPn0036	51002	51796	F	CT338 hypothetical protein
CPn0037	51792	52115	F	ptsH-PTS Phosphocarrier Protein Hpr-(CT337)
CPn0038	52119	53831	F	ptsI-PTS PEP Phosphotransferase-(CT336)
CPn0039	54250	53963	R	ybaB-(CT335)
CPn0040	55643	54318	R	dnaX_1-DNA Pol III Gamma and Tau_1-(CT334)
CPn0041	55996	57342	F	
CPn0042	57403	58182	F	
CPn0043	58447	60372	F	
CPn0044	60419	60778	F	
CPn0045	61069	62790	F	
CPn0046	62790	63263	F	
CPn0047	63455	63652	F	
CPn0048	63687	65801	F	*yqfF-Bs conserved hypothetical IM protein
CPn0049	66296	65817	R	
CPn0050	66813	66499	R	
CPn0051	66833	67111	F	
CPn0052	68005	67304	R	hemC-Porphobilinogen Deaminase-(CT299)
CPn0053	69344	67986	R	sms-Sms Protein-(CT298)
CPn0054	70023	69313	R	rnc-Ribonuclease III-(CT297)
CPn0055	70129	70590	F	CT296 hypothetical protein
CPn0056	70953	72746	F	mrsA-Phosphomannomutase-(CT295)
CPn0057	72934	73554	F	sodM-Superoxide Dismutase (Mn)-(CT294)
CPn0058	73639	74562	F	accD-AcCoA Carboxylase/Transferase Beta-(CT293)
CPn0059	74616	75050	F	dut-dUTP Nucleotidohydrolase-(CT292)
CPn0060	75055	75528	F	ptsN_1-PTS IIA Protein-(CT291)
CPn0061	75534	76208	F	ptsN_2-PTS IIA Protein + HTH DNA-Binding Domain-(CT290)
CPn0062	76308	77690	F	CT289 hypothetical protein
CPn0063	78112	78267	F	
CPn0064	78346	78576	F	
CPn0065	78924	80651	F	CT288 hypothetical protein
CPn0066	80925	82655	F	

CPn0067	82953	84053	F	
CPn0068	84903	84331	R	CT360 hypothetical protein
CPn0069	85236	87086	F	
CPn0070	87378	87208	R	
CPn0071	88045	87599	R	CT325 hypothetical protein
CPn0072	89061	88057	R	CT324 hypothetical protein
CPn0073	89356	89574	F	infA-Initiation Factor IF-1-(CT323)
CPn0074	89774	90955	F	tufA-Elongation Factor Tu-(CT322)
CPn0075	91102	91350	F	secE-preprotein translocase-(CT321)
CPn0076	91358	91903	F	nusG-Transcriptional Antitermination-(CT320)
CPn0077	92013	92435	F	rl11-L11 Ribosomal Protein-(CT319)
CPn0078	92465	93160	F	rl1-L1 Ribosomal Protein-(CT318)
CPn0079	93179	93688	F	rl10-L10 Ribosomal Protein-(CT317)
CPn0080	93735	94121	F	rl7-L7/L12 Ribosomal Protein-(CT316)
CPn0081	94261	98016	F	rpoB-RNA Polymerase Beta-(CT315)
CPn0082	98043	102221	F	rpoC-RNA Polymerase Beta'-(CT314)
CPn0083	102332	103312	F	tal-Transaldolase-(CT313)
CPn0084	103362	103751	F	predicted ferredoxin-(CT312)
CPn0085	104506	103766	R	CT311 hypothetical protein
CPn0086	104904	105527	F	atpE-ATP Synthase Subunit E-(CT310)
CPn0087	105579	106376	F	CT309 hypothetical protein
CPn0088	106373	108145	F	atpA-ATP Synthase Subunit A-(CT308)
CPn0089	108153	109466	F	atpB-ATP Synthase Subunit B-(CT307)
CPn0090	109454	110080	F	atpD-ATP Synthase Subunit D-(CT306)
CPn0091	110074	112053	F	atpI-ATP Synthase Subunit I-(CT305)
CPn0092	112151	112573	F	atpK-ATP Synthase Subunit K-(CT304)
CPn0093	112509	113015	F	CT303 hypothetical protein
CPn0094	113152	115971	F	valS-Valyl tRNA Synthetase-(CT302)
CPn0095	116037	118790	F	pknD-S/T Protein Kinase-(CT301)
CPn0096	124314	118837	R	uvrA-Excinuclease ABC Subunit A-(CT333)
CPn0097	124555	126006	F	pyk-Pyruvate Kinase-(CT332)
CPn0098	127491	126091	R	htrB-Acyltransferase-(CT010)
CPn0099	127593	127865	F	
CPn0100	129141	127882	R	CT011 hypothetical protein
CPn0101	129932	129141	R	ybbP family hypothetical protein-(CT012)
CPn0102	130123	131466	F	cydA-Cytochrome Oxidase Subunit I-(CT013)
CPn0103	131480	132511	F	cydB-Cytochrome Oxidase Subunit II-(CT014)
CPn0104	133875	132676	R	CT017 hypothetical protein
CPn0105	134847	134029	R	CT016 hypothetical protein
CPn0106	135091	136374	F	phoH-ATPase-(CT015)
CPn0107	137162	136392	R	CT058 hypothetical protein_1
CPn0108	137857	137303	R	CT018
CPn0109	138655	141783	F	ileS-Isoleucyl-tRNA Synthetase-(CT019)
CPn0110	143734	141827	R	lepB-Signal Peptidase I-(CT020)
CPn0111	144686	143934	R	CT021 hypothetical protein
CPn0112	144767	145093	F	rl31-L31 Ribosomal Protein-(CT022)
CPn0113	145335	146405	F	pfrA-Peptide Chain Releasing Factor (RF-1)-(CT023)
CPn0114	146398	147261	F	hemK-A/G specific methylase-(CT024)
CPn0115	147279	148622	F	ffh-Signal Recognition Particle GTPase-(CT025)
CPn0116	148616	148972	F	rs16-S16 Ribosomal Protein-(CT026)
CPn0117	148989	150071	F	trmD-tRNA (guanine N-1)-Methyltransferase-(CT027)
CPn0118	150102	150464	F	rl19-L19 Ribosomal Protein-(CT028)
CPn0119	150523	151164	F	rnhB_1-Ribonuclease HII_1-(CT029)
CPn0120	151164	151778	F	gmK-GMP Kinase-(CT030)
CPn0121	151778	152068	F	CT031 hypothetical protein
CPn0122	152071	153723	F	metG-Methionyl-tRNA Synthetase-(CT032)
CPn0123	155969	153774	R	recD_1-Exodeoxyribonuclease V (Alpha Subunit)_1-(CT033)
CPn0124	156614	158068	F	
CPn0125	158096	158605	F	
CPn0126	158809	161085	F	
CPn0127	162143	161130	R	ytff-Cationic Amino Acid Transporter-(CT034)
CPn0128	162277	163053	F	bpl1-Biotin Protein Ligase-(CT035)
CPn0129	163717	163064	R	similarity to CT036
CPn0130	164245	163751	R	
CPn0131	164549	165580	F	
CPn0132	165587	166561	F	
CPn0133	167334	166564	R	CHLPS hypothetical protein-(CT109)
CPn0134	169098	167467	R	groEL_1-HSP-60_1-(CT110)
CPn0135	169448	169143	R	groES-10KDa Chaperonin-(CT111)
CPn0136	171401	169569	R	pepF-Oligopeptidase-(CT112)
CPn0137	172254	171502	R	ybgI-ACR family-(CT108)
CPn0138	174019	172700	R	hemL-Glutamate-1-semialdehyde-2,1-aminomutase-(CT210)

CPn0139	174656	174093	R	yqgE-(CT210)
CPn0140	175110	174673	R	yqdE-(CT212)
CPn0141	175802	175110	R	rpiA-Ribose-5-P Isomerase A-(CT213)
CPn0142	176091	175816	R	
CPn0143	177335	176214	R	*yxjG_Bs_1 Hypothetical Protein
CPn0144	177963	180560	F	clpB-Clp Protease ATPase-(CT113)
CPn0145	180777	182369	F	CT114 hypothetical protein
CPn0146	182613	183095	F	
CPn0147	183225	183671	F	
CPn0148	183846	185702	F	pknI-S/T Protein Kinase-(CT145)
CPn0149	185715	187700	F	dnlJ-DNA Ligase-(CT146)
CPn0150	187834	192444	F	CT147 hypothetical protein
CPn0151	194142	192625	R	mhpA-Monooxygenase-(CT148)
CPn0152	195265	194318	R	CT149 hypothetical protein
CPn0153	195433	197892	F	leuS-Leucyl tRNA Synthetase-(CT209)
CPn0154	197892	199202	F	gseA-KDO Transferase-(CT208)
CPn0155	199691	199488	R	
CPn0156	200117	199770	R	
CPn0157	200723	200298	R	
CPn0158	201430	200894	R	
CPn0159	201772	201467	R	
CPn0160	203791	202127	R	pfkA_1-Fructose-6-P Phosphotransferase_1-(CT207)
CPn0161	204622	203798	R	predicted acyltransferase family-(CT206)
CPn0162	205828	204803	R	
CPn0163	206026	206394	F	
CPn0164	206498	206998	F	
CPn0165	206998	207582	F	
CPn0166	207630	207962	F	
CPn0167	208306	207977	R	
CPn0168	208641	208417	R	
CPn0169	209501	208710	R	
CPn0170	211026	210025	R	
CPn0171	212435	211149	R	*guaA-GMP Synthase
CPn0172	213177	212440	R	*guaB/impD-Inosine 5'-monophosphate dehydrogenase (C00H-terminal region only)
CPn0173	213987	213715	R	
CPn0174	214257	214724	F	
CPn0175	214898	215275	F	
CPn0176	215286	216518	F	CT153 hypothetical protein
CPn0177	217459	216608	R	
CPn0178	218052	217789	R	
CPn0179	218403	218056	R	
CPn0180	218851	218355	R	
CPn0181	219175	218777	R	
CPn0182	220695	219334	R	accC-Biotin Carboxylase-(CT124)
CPn0183	221195	220695	R	accB-Biotin Carboxyl Carrier Protein-(CT123)
CPn0184	221775	221221	R	efp_1-Elongation Factor P_1-(CT122)
CPn0185	222451	221765	R	rpe/araD-Ribulose-P Epimerase-(CT121)
CPn0186	222899	224068	F	*similarity to Cps Inca_1-(CT119)
CPn0187	224248	225045	F	predicted methylase-(CT133)
CPn0188	225111	226400	F	CT132 hypothetical protein
CPn0189	226400	229825	F	CT131 homolog-(Possible Transmembrane Protein)
CPn0190	229919	231274	F	
CPn0191	231991	231314	R	glnQ-ABC Amino Acid Transporter ATPase-(CT130)
CPn0192	232634	231984	R	glnP-ABC Amino Acid Transporter Permease-(CT129)
CPn0193	233126	232686	R	*argR-Arginine Repressor
CPn0194	233210	234241	F	gcp_2-O-Sialoglycoprotein Endopeptidase_2-(CT197)
CPn0195	234190	235785	F	oppA_1-Oligopeptide Binding Protein_1
CPn0196	235939	237519	F	oppA_2-Oligopeptide Binding Protein_2-(CT198)
CPn0197	237578	238882	F	oppA_3-Oligopeptide Binding Protein_3
CPn0198	239169	240746	F	oppA_4-Oligopeptide Binding Protein_4
CPn0199	241042	241983	F	oppB_1-Oligopeptide Permease_1-(CT199)
CPn0200	242017	242868	F	oppC_1-Oligopeptide Permease_1-(CT200)
CPn0201	242864	243715	F	oppD-Oligopeptide Transport ATPase-(CT201)
CPn0202	243715	244500	F	oppF-Oligopeptide Transport ATPase-(CT202)
CPn0203	245008	245802	F	
CPn0204	245817	246002	F	
CPn0205	246133	246327	F	
CPn0206	246409	247161	F	CT203 hypothetical protein
CPn0207	247208	248617	F	ybhI/sodiT1-Oxoglutarate/Malate Translocator-(CT204)
CPn0208	248953	250602	F	pfkA_2-Fructose-6-P Phosphotransferase_2-(CT205)
CPn0209	251036	251272	F	

CPn0210	252384	251440	R	
CPn0211	252756	252463	R	
CPn0212	254066	252888	R	
CPn0213	254342	254190	R	
CPn0214	255657	254446	R	
CPn0215	257015	255759	R	
CPn0216	257608	257174	R	
CPn0217	257896	258579	F	ypdP-(CT140)
CPn0218	259058	258582	R	
CPn0219	259357	260472	F	tgt-Queuine tRNA Ribosyl Transferase-(CT193)
CPn0220	260696	261238	F	
CPn0221	261657	262064	F	
CPn0222	262504	262842	F	*weak similarity to Bacteriophage CHP1 (Orf4)
CPn0223	262956	263333	F	
CPn0224	263435	263674	F	
CPn0225	263873	264541	F	
CPn0226	264566	264967	F	
CPn0227	265416	265009	R	dsbB-Disulfide bond Oxidoreductase-(CT176)
CPn0228	266110	265412	R	dsbG-Disulfide Bond Chaperone-(CT177)
CPn0229	266328	267560	F	CT178 hypothetical protein
CPn0230	268253	267576	R	CT179 hypothetical protein
CPn0231	268957	268253	R	tauB-ABC Transport ATPase (Nitrate/Fe)-(CT180)
CPn0232	270122	269232	R	*similarity to 5'-Methylthioadenosine / S-Adenosylhomocysteine Nucleosidase
CPn0233	270424	270248	R	
CPn0234	271240	270548	R	CT181 hypothetical protein
CPn0235	271416	272177	F	kdsB-deoxyoctulononic Acid Synthetase-(CT182)
CPn0236	272156	273766	F	pyrG-CTP Synthetase-(CT183)
CPn0237	273762	274214	F	yggF Family-(CT184)
CPn0238	274303	275838	F	zwf-Glucose-6-P Dehydrogenase-(CT185)
CPn0239	275899	276672	F	devB-Glucose-6-P Dehydrogenase (DevB family)-(CT186)
CPn0240	277861	276698	R	
CPn0241	279354	278203	R	
CPn0242	279918	279487	R	
CPn0243	280555	280133	R	
CPn0244	280918	281556	F	adk-Adenylate Kinase-(CT128)
CPn0245	281645	282499	F	ydH-Polysaccharide Hydrolase-Invasin Repeat Family-(CT127)
CPn0246	282952	282551	R	rs9-S9 Ribosomal Protein-(CT126)
CPn0247	283415	282969	R	rl13-L13 Ribosomal Protein-(CT125)
CPn0248	284327	283650	R	ycfV/ybbA-ABC Transporter ATPase-(CT152)
CPn0249	285841	284333	R	CT151 hypothetical protein
CPn0250	286057	285902	R	rl33-L33 Ribosomal Protein-(CT150)
CPn0251	286060	287559	F	*conserved hypothetical protein
CPn0252	288112	287576	R	CT144 hypothetical protein (frame-shift with 0253?)
CPn0253	288456	287950	R	CT144 hypothetical protein_1
CPn0254	289262	288459	R	CT143 hypothetical protein_1
CPn0255	290165	289329	R	CT142 hypothetical protein_1
CPn0256	291264	290398	R	CT144 hypothetical protein_2
CPn0257	292127	291267	R	CT143 hypothetical protein_2
CPn0258	292534	292133	R	CT142 hypothetical protein (frame-shift with 0259?)
CPn0259	292986	292441	R	CT142 hypothetical protein_2
CPn0260	294045	293548	R	secA_1-Protein Translocase Subunit_1-(CT141)
CPn0261	294302	295033	F	ydaO-PP-Loop Superfamily ATPase-(CT217)
CPn0262	295091	295933	F	surE-SurE-like Acid Phosphatase-(CT218)
CPn0263	296249	297136	F	yqfU hypothetical protein-(CT221)
CPn0264	297730	297155	R	ubiD-Phenylacrylate Decarboxylase-(CT220)
CPn0265	298620	297730	R	ubiA-Benzoate Octaphenyltransferase-(CT219)
CPn0266	299184	299876	F	
CPn0267	300122	300910	F	
CPn0268	300935	301318	F	
CPn0269	302450	301476	R	Dipeptidase-(CT138)
CPn0270	303325	302468	R	ywlC-SuA5 Superfamily-related Protein-(CT137)
CPn0271	303634	304362	F	Lysophospholipase esterase-(CT136)
CPn0272	305233	304340	R	dnaX_2-DNA Pol III Gamma and Tau_2-(CT187)
CPn0273	305844	305227	R	tdk-Thymidylate Kinase-(CT188)
CPn0274	308353	305852	R	gyrA_1-DNA Gyrase Subunit A_1-(CT189)
CPn0275	310786	308372	R	gyrB_1-DNA Gyrase Subunit B_1-(CT190)
CPn0276	311137	310793	R	CT191 hypothetical protein
CPn0277	311910	311404	R	
CPn0278	312875	312060	R	*conserved outer membrane lipoprotein protein
CPn0279	313537	312875	R	*Possible ABC Transporter Permease Protein
CPn0280	314572	313550	R	dppF-Dipeptide Transporter ATPase-(CT689)

CPn0281	315057	316103	F	dhnA-Predicted 1,6-Fructose Biphosphate Aldolase (dehydrin family)-(CT215)
CPn0282	316126	317529	F	xasA/gadC-Amino Acid Transporter-(CT216)
CPn0283	318497	317532	R	
CPn0284	319045	318551	R	
CPn0285	320595	319051	R	
CPn0286	322059	320650	R	mgtE-Mg++ Transporter (CBS Domain)-(CT194)
CPn0287	324221	322089	R	
CPn0288	325716	324571	R	CT195 hypothetical protein
CPn0289	325812	326996	F	aaaT-Neutral Amino Acid (Glutamate) Transporter-(CT230)
CPn0290	327042	328523	F	Na-dependent Transporter-(CT231)
CPn0291	328667	329194	F	incB-Inclusion Membrane Protein B-(CT232)
CPn0292	329228	329836	F	incC-Inclusion Membrane Protein C-(CT233)
CPn0293	329949	332723	F	CT234 hypothetical protein
CPn0294	333092	333502	F	cAMP-Dependent Protein Kinase Regulatory Subunit-(CT235)
CPn0295	333863	333627	R	acpP-Acyl Carrier Protein-(CT236)
CPn0296	334765	334022	R	fabG-Oxoacyl (Carrier Protein) Reductase-(CT237)
CPn0297	335697	334774	R	fabD-Malonyl Acyl Carrier Transcyclase-(CT238)
CPn0298	336721	335717	R	fabH-Oxoacyl Carrier Protein Synthase III-(CT239)
CPn0299	336816	337415	F	recR-Recombination Protein-(CT240)
CPn0300	337783	340152	F	yaeT-Omp85 Analog-(CT241)
CPn0301	340250	340762	F	(OmpH-Like Outer Membrane Protein)-(CT242)
CPn0302	340787	341866	F	lpxD-UDP Glucosamine N-Acyltransferase-(CT243)
CPn0303	342958	341921	F	CT244 hypothetical protein
CPn0304	343133	344158	F	pdhA/odpA-Pyruvate Dehydrogenase Alpha-(CT245)
CPn0305	344154	345137	F	pdhB/odpB-Pyruvate Dehydrogenase Beta-(CT246)
CPn0306	345145	346431	F	pdhC-Dihydrolipoamide Acetyltransferase-(CT247)
CPn0307	348986	346515	F	glgP-Glycogen Phosphorylase-(CT248)
CPn0308	349234	349596	F	similarity to CT249
CPn0309	350974	349595	R	dnaA_1-Replication Initiation Protein_1-(CT250)
CPn0310	353433	351049	R	60IM-60kDa Inner Membrane Protein-(CT251)
CPn0311	354438	353575	R	lgt-Prolipoprotein Diacylglycerol Transferase-(CT252)
CPn0312	354524	354976	F	CT101 hypothetical protein
CPn0313	354990	355355	F	acpS-Acyl-carrier Protein Synthase-(CT100)
CPn0314	356285	355353	R	trxB-Thioredoxin Reductase-(CT099)
CPn0315	356977	358716	F	rs1-S1 Ribosomal Protein-(CT098)
CPn0316	358820	360121	F	nusA-N Utilization Protein A-(CT097)
CPn0317	360081	362750	F	infB-Initiation Factor-2-(CT096)
CPn0318	362767	363126	F	rbfA-Ribosome Binding Factor A-(CT095)
CPn0319	363175	363879	F	truB-tRNA Pseudouridine Synthase-(CT094)
CPn0320	363860	364783	F	ribF-FAD Synthase-(CT093)
CPn0321	365858	364767	R	yehF-GTP Binding Protein-(CT092)
CPn0322	366249	367328	F	yscU-YopS Translocation Protein U -(CT091)
CPn0323	367331	369460	F	lcrD- Low Calcium Response D-(CT090)
CPn0324	369492	370688	F	lcrE- Low Calcium Response E-(CT089)
CPn0325	370708	371148	F	syce-Secretion Chaperone-(CT088)
CPn0326	371148	372725	F	malQ-Glucanotransferase-(CT087)
CPn0327	372945	373211	F	rl28-L28 Ribosomal Protein-(CT086)
CPn0328	373241	374992	F	CT085 hypothetical protein
CPn0329	375088	376146	F	Phospholipase D Superfamily [leader (33) peptide]-(CT084)
CPn0330	376675	376202	R	CT083 hypothetical protein
CPn0331	378437	376701	R	CT082 hypothetical protein
CPn0332	378655	378536	R	CHLTR T2 Protein-(CT081)
CPn0333	379090	378800	R	ltuB-(CT080)
CPn0334	379311	379823	F	CT079 similarity
CPn0335	379817	380674	F	fold-Methylene Tetrahydrofolate Dehydrogenase-(CT078)
CPn0336	380650	381591	F	yoyL-(CT077)
CPn0337	382027	381575	R	smpB- Small Protein B-(CT076)
CPn0338	382278	383375	F	dnaN-DNA Pol III (beta chain)-(CT075)
CPn0339	383420	384034	F	recF-ABC superfamily ATPase-(CT074)
CPn0340	383842	384156	F	(frame-shift with 0339)
CPn0341	384160	384495	F	(frame-shift with 0340)
CPn0342	384622	385062	F	predicted OMP [leader (19) peptide]-(CT073)
CPn0343	384999	385595	F	(frame-shift with 0342?)
CPn0344	387420	385558	R	yaeL-Metalloprotease-(CT072)
CPn0345	388572	387436	R	yaeM-(CT071)
CPn0346	389675	388704	R	troD/ytgD-Integral Membrane Protein-(CT070)
CPn0347	391021	389678	R	troC/ytgC-Integral Membrane Protein-(CT069)
CPn0348	391803	391027	R	troB/ytgB-ABC transporter ATPase-(CT068)
CPn0349	392770	391790	R	troA/ytgA-Solute Protein Binding Family-(CT067)
CPn0350	393181	393684	F	CT066 hypothetical protein
CPn0351	393888	395432	F	adt_1-ADP/ATP Translocase_1-(CT065)

CPn0352	395574	396830	F	
CPn0353	396893	397135	F	
CPn0354	397167	398507	F	
CPn0355	399889	398591	R	
CPn0356	400459	400109	R	
CPn0357	401317	400469	R	
CPn0358	401751	401578	R	
CPn0359	402012	403817	F	lepA-GTPase-(CT064)
CPn0360	405358	403922	R	gnd-6-Phosphogluconate Dehydrogenase-(CT063)
CPn0361	406647	405382	R	tyrS-tyrosyl tRNA Synthetase-(CT062)
CPn0362	407825	407055	R	fliA/rpsD-Sigma-28/WhiG Family-(CT061)
CPn0363	409688	407943	R	flhA-Flagellar Secretion Protein-(CT060)
CPn0364	409966	410238	F	fer4-Ferredoxin IV-(CT059)
CPn0365	410528	411544	F	
CPn0366	411976	412440	F	
CPn0367	413102	413836	F	
CPn0368	413790	414107	F	
CPn0369	414351	415562	F	CT058 hypothetical protein_2
CPn0370	415800	416912	F	CT058 hypothetical protein_3
CPn0371	417147	417503	F	
CPn0372	417687	418001	F	
CPn0373	418380	420218	F	gcpE-(CT057)
CPn0374	420218	420961	F	CT056 hypothetical protein
CPn0375	421121	421615	F	
CPn0376	421854	422294	F	
CPn0377	423438	422347	R	sucB_1-Dihydrolipoamide Succinyltransferase_1-(CT055)
CPn0378	426168	423445	R	sucA-Oxoglutarate Dehydrogenase-(CT054)
CPn0379	426322	426765	F	CT053 hypothetical protein
CPn0380	426758	427876	F	hemN_1-Coproporphyrinogen III Oxidase_1-(CT052)
CPn0381	429809	428037	R	CT326 similarity
CPn0382	430749	430036	R	yabC/yraL-SAM-Dependent Methytransferase-(CT048)
CPn0383	431693	430749	R	CT047 hypothetical protein
CPn0384	432377	431862	R	hctB-Histone-like Protein 2-(CT046)
CPn0385	434018	432522	R	pepA-Leucyl Aminopeptidase A-(CT045)
CPn0386	434525	434046	R	ssb-SS DNA Binding Protein-(CT044)
CPn0387	435196	434699	R	CT043 hypothetical protein
CPn0388	435329	437320	F	glgX-Glycogen Hydrolase (debranching)-(CT042)
CPn0389	438134	437319	R	CT041 hypothetical protein
CPn0390	439144	438134	R	ruvB-Holliday Junction Helicase-(CT040)
CPn0391	439692	439510	R	
CPn0392	439814	440383	F	dcd-dCTP Deaminase-(CT039)
CPn0393	440379	440723	F	CT038 hypothetical protein
CPn0394	440736	441968	F	tlyC_1-CBS Domain protein (Hemolysin Homolog)_1-(CT256)
CPn0395	441964	443175	F	CT257 hypothetical protein
CPn0396	444353	443241	R	yhfo-Nifs-related protein-(CT258)
CPn0397	445115	444381	R	PP2C phosphatase family-(CT259)
CPn0398	445533	445700	F	
CPn0399	445879	446523	F	CT253 hypothetical protein
CPn0400	446536	447306	F	CT254 hypothetical protein
CPn0401	447884	447495	R	CT255 hypothetical protein
CPn0402	448994	447888	R	mutY-Adenine Glycosylase-(CT107)
CPn0403	449015	449710	F	yceC-predicted pseudouridine synthetase family-(CT106)
CPn0404	450887	449871	R	
CPn0405	451739	450966	R	CT105 hypothetical protein
CPn0406	451969	452865	F	fabI-Enoyl-Acyl-Carrier Protein Reductase-(CT104)
CPn0407	453742	452858	R	HAD superfamily hydrolase/phosphatase-(CT103)
CPn0408	454105	454581	F	CT102 hypothetical protein
CPn0409	454645	455127	F	CT260 hypothetical protein
CPn0410	455123	455833	F	dnaQ_1-DNA Pol III Epsilon Chain_1-(CT261)
CPn0411	455833	456609	F	CT262 hypothetical protein
CPn0412	456590	457246	F	CT263 hypothetical protein
CPn0413	459203	457227	R	msbA-Transport ATP Binding Protein-(CT264)
CPn0414	460143	459172	R	accA-AcCoA Carboxylase/Transferase Alpha-(CT265)
CPn0415	461498	460221	R	CT266 hypothetical protein
CPn0416	461856	461557	R	himD/ihfA-Integration Host Factor Alpha-(CT267)
CPn0417	463035	462244	R	amiA-N-Acetylmuramoyl Alanine Amidase-(CT268)
CPn0418	464401	462953	R	murE-N-Acetylmuramoylalanylglutamyl DAP Ligase-(CT269)
CPn0419	466834	464876	R	pbp3- transglycolase/transpeptidase-(CT270)
CPn0420	467108	466824	R	CT271 hypothetical protein
CPn0421	467998	467108	R	yabC-PBP2B Family methyltransferase-(CT272)
CPn0422	468242	468784	F	CT273 hypothetical protein
CPn0423	468791	469216	F	CT274 hypothetical protein

CPn0424	469612	470961	F	dnaA_2-Replication Initiation Factor_4-(CT275)
CPn0425	470980	471564	F	CT276 hypothetical proteins
CPn0426	472111	471536	R	CT277 similarity
CPn0427	472207	473715	F	nqr2-NADH (Ubiquinone) Dehydrogenase-(CT278)
CPn0428	473722	474681	F	nqr3-NADH (Ubiquinone) Oxidoreductase, Gamma-(CT279)
CPn0429	474681	475319	F	nqr4-NADH (Ubiquinone) Reductase 4-(CT280)
CPn0430	475326	476093	F	nqr5-NADH (Ubiquinone) Reductase 5-(CT281)
CPn0431	476483	476151	R	
CPn0432	476816	476514	R	
CPn0433	477273	476929	R	gcsH-Glycine Cleavage System H Protein-(CT282)
CPn0434	479462	477276	R	CT283 hypothetical protein
CPn0435	480902	479475	R	Phospholipase D superfamily [uncleavable leader peptide]-(CT284)
CPn0436	481618	480902	R	lplA-Lipoate Protein Ligase-Like Protein-(CT285)
CPn0437	481816	484350	F	clpC-ClpC Protease-(CT286)
CPn0438	485416	484334	R	ycbF-PP-loop superfamily ATPase-(CT287)
CPn0439	485553	486077	F	
CPn0440	486105	486740	F	
CPn0441	486891	487838	F	CT007 hypothetical protein
CPn0442	488013	488528	F	CT006 hypothetical protein
CPn0443	488729	489979	F	CT005 hypothetical protein
CPn0444	490287	494507	F	pmp_6-Polymorphic Outer Membrane Protein G/I Family
CPn0445	494772	497579	F	pmp_7-Polymorphic Outer Membrane Protein G Family
CPn0446	497626	500415	F	pmp_8-Polymorphic Outer Membrane Protein G Family
CPn0447	500568	503351	F	pmp_9-Polymorphic Outer Membrane Protein G/I Family
CPn0448	504810	503698	R	*yxjG_Bs_2 Hypothetical Protein
CPn0449	507231	505330	R	pmp_10-PMP_10 (Frame-shift with 0451)
CPn0450	508112	507180	R	pmp_10-Polymorphic Outer Membrane Protein G Family
CPn0451	508275	511058	F	pmp_11-Polymorphic Outer Membrane Protein G Family
CPn0452	511319	512860	F	pmp_12-Polymorphic Outer Membrane Protein A/I Family (truncated)
CPn0453	513234	516152	F	pmp_13 -Polymorphic Outer Membrane Protein G Family
CPn0454	516182	519115	F	pmp_14-Polymorphic Outer Membrane Protein H Family
CPn0455	520348	519458	R	
CPn0456	521532	520327	R	
CPn0457	523865	522120	R	
CPn0458	526320	524236	R	
CPn0459	527005	526619	R	
CPn0460	527840	526992	R	
CPn0461	528638	527844	R	
CPn0462	531052	529037	R	
CPn0463	532357	531191	R	
CPn0464	532842	532366	R	
CPn0465	533212	532871	R	
CPn0466	533724	536537	F	pmp_15-Polymorphic Outer Membrane Protein E Family
CPn0467	536633	539434	F	pmp_16-Polymorphic Outer Membrane Protein E Family
CPn0468	539632	540432	F	pmp_17-Polymorphic Outer Membrane Protein E Family
CPn0469	540399	541460	F	pmp_17-Polymorphic Outer Membrane Protein (Frame-shift with 0469)
CPn0470	541357	542532	F	pmp_17-Polymorphic Outer Membrane Protein (Frame-shift with 0470)
CPn0471	542564	545401	F	pmp_18-Polymorphic Outer Membrane Protein E/F Family
CPn0472	547905	545581	R	
CPn0473	549593	548070	R	
CPn0474	551573	549807	R	CT365 hypothetical protein
CPn0475	553844	551685	R	glgB-Glucan Branching Enzyme-(CT866)
CPn0476	554844	553858	R	CT865 hypothetical protein
CPn0477	556106	554844	R	*yqeV_Bs Hypothetical Protein
CPn0478	557625	556210	R	hflX-GTP Binding Protein-(CT379)
CPn0479	558425	557616	R	phnP-Metal Dependent Hydrolase-(CT380)
CPn0480	559303	558650	R	CT383 hypothetical protein
CPn0481	560946	559339	R	
CPn0482	561737	560961	R	artJ-Arginine Periplasmic Binding Protein-(CT381)
CPn0483	561836	564964	F	
CPn0484	564970	565824	F	aroG-Deoxyheptonate Aldolase-(CT382)
CPn0485	566038	566229	F	CT382.1 hypothetical protein
CPn0486	567784	566405	R	*hypothetical proline permease
CPn0487	569740	568112	R	CT384 hypothetical protein
CPn0488	570096	569767	R	hitA-HIT Family Hydrolase-(CT385)
CPn0489	570965	570096	R	CT386 hypothetical protein
CPn0490	571279	573333	F	CT387 hypothetical protein
CPn0491	574352	573336	R	CT389 hypothetical protein
CPn0492	574652	574804	F	
CPn0493	575004	574855	R	
CPn0494	575364	575146	R	
CPn0495	575603	576793	F	aspC-Aspartate Aminotransferase-(CT390)

CPn0496	576793	577812	F	CT391 hypothetical protein
CPn0497	578089	577820	R	CT388 hypothetical protein
CPn0498	579035	578085	R	
CPn0499	580359	579205	R	
CPn0500	580659	582362	F	proS-Prolyl tRNA Synthetase-(CT393)
CPn0501	582457	583650	F	hrcA-HTH Transcriptional Repressor-(CT394)
CPn0502	583650	584201	F	grpE-HSP-70 Cofactor-(CT395)
CPn0503	584234	586213	F	dnaK-HSP-70-(CT396)
CPn0504	586487	588514	F	vacB-ribonuclease family-(CT397)
CPn0505	588519	589106	F	*3-methyladenine DNA glycosylase
CPn0506	589172	589840	F	CT421 hypothetical protein
CPn0507	589961	590122	F	CT421.1 hypothetical protein
CPn0508	590142	590300	F	CT421.2 hypothetical protein
CPn0509	590335	590808	F	(predicted Metalloenzyme)-(CT422)
CPn0510	590813	591973	F	tlyC_2-CBS Domains (Hemolysin homolog)_2-(CT423)
CPn0511	592141	592488	F	rsbV_1-Sigma Regulatory Factor_1-(CT424)
CPn0512	592553	594412	F	CT425 hypothetical protein
CPn0513	594647	595753	F	Fe-S oxidoreductase_1-(CT426)
CPn0514	595729	596520	F	CT427 hypothetical protein
CPn0515	596492	597181	F	ubiE-Ubiquinone Methyltransferase-(CT428)
CPn0516	598814	597255	R	
CPn0517	599631	598795	R	
CPn0518	600803	599832	R	CT429 hypothetical protein
CPn0519	601674	600904	R	dapF-Diaminopimelate Epimerase-(CT430)
CPn0520	602218	601646	R	clpP-CLP Protease-(CT431)
CPn0521	603797	602241	R	glyA-Serine Hydroxymethyltransferase-(CT432)
CPn0522	603987	604655	F	CT433 hypothetical protein
CPn0523	604723	605052	F	
CPn0524	605103	606179	F	
CPn0525	606522	607283	F	CT398 hypothetical protein
CPn0526	608696	607710	R	yrbH-GutQ/KpsF Family Sugar-P Isomerase-(CT399)
CPn0527	609904	608726	R	sucB_2-Dihydrolipoamide Succinyltransferase_2-(CT400)
CPn0528	611162	609921	R	gltT-Glutamate Symport-(CT401)
CPn0529	612259	611165	R	ycaH-ATPase-(CT402)
CPn0530	613254	612460	R	spoU_1-rRNA Methylase_1-(CT403)
CPn0531	614069	613245	R	SAM dependent methyltransferase-(CT404)
CPn0532	614674	614075	R	ribC/risA-Riboflavin Synthase-(CT405)
CPn0533	614930	615385	F	CT406 hypothetical protein
CPn0534	615413	615784	F	dksA-DnaK Suppressor-(CT407)
CPn0535	615793	616296	F	lspA-Lipoprotein Signal Peptidase-(CT408)
CPn0536	616345	617691	F	dagA_1-D-Ala/Gly Permease_1-(CT409)
CPn0537	617833	618189	F	CT814.1 hypothetical protein
CPn0538	618212	618511	F	CT814 hypothetical protein
CPn0539	618705	621545	F	pmp_19-polymorphic outer membrane protein A Family -(CT412)
CPn0540	621694	626862	F	pmp_20-polymorphic outer membrane protein B Family-(CT413)
CPn0541	627170	628003	F	Solute binding protein (-yebL-Synechocystis Adhesin Homolog)-(CT415)
CPn0542	628003	628737	F	ABC Transporter ATPase-(CT416)
CPn0543	628725	629603	F	(Metal Transport Protein)-(CT417)
CPn0544	630529	629525	R	yhbZ-GTP binding protein-(CT418)
CPn0545	630884	630633	R	rl27-L27 ribosomal protein-(CT419)
CPn0546	631229	630912	R	rl21-L21 Ribosomal Protein-(CT420)
CPn0547	631661	632188	F	ygbB family-(CT434)
CPn0548	633231	632191	R	cysJ-Sulfite Reductase-(CT435)
CPn0549	633569	633255	R	rs10-S10 Ribosomal Protein-(CT436)
CPn0550	635661	633580	R	fusA-Elongation Factor G-(CT437)
CPn0551	636168	635698	R	rs7-S7 Ribosomal Protein-(CT438)
CPn0552	636587	636219	R	rs12-S12 Ribosomal Protein-(CT439)
CPn0553	637747	636812	R	
CPn0554	637854	638141	F	CT440 hypothetical protein
CPn0555	638298	640241	F	tsp-Tail-Specific Protease-(CT441)
CPn0556	640912	640325	R	crpA-15kDa Cysteine-Rich Protein-(CT442)
CPn0557	642861	641194	R	omcB-60kDa Cysteine-Rich Outer Membrane Complex Protein-(CT443)
CPn0558	643300	643031	R	omcA-9kDa-Cysteine-Rich Outer Membrane Complex Lipoprotein-(CT444)
CPn0559	643742	643927	F	CT441.1 hypothetical protein
CPn0560	645612	644098	R	gltX-Glutamyl-tRNA Synthetase-(CT445)
CPn0561	646404	645871	R	euo-CHLPS Euo Protein-(CT446)
CPn0562	648036	646918	R	*CHLPS 43 kDa protein homolog_1
CPn0563	650056	648293	R	recJ-ssDNA Exonuclease-(CT447)
CPn0564	654350	650145	R	secD&secF-Protein Export Proteins SecD/SecF (fusion)-(CT448)
CPn0565	655630	654533	R	CT449 hypothetical protein
CPn0566	656141	656890	F	yaeS family-(CT450)
CPn0567	656894	657817	F	cdsA-Phosphatidate Cytidylyltransferase-(CT451)

CPn0568	657817	658464	F	cdsA-Phosphatidate Cytidylyltransferase-(CT452)
CPn0569	658464	659099	F	plsC-Glycerol-3-P Acyltransferase-(CT453)
CPn0570	659107	660789	F	argS-Arginyl tRNA Transferase-(CT454)
CPn0571	662122	660749	R	murA-UDP-N-Acetylglucosamine Transferase-(CT455)
CPn0572	662352	664616	F	CT456 hypothetical protein
CPn0573	665404	664691	R	yebC family-(CT457)
CPn0574	665945	665394	R	
CPn0575	666494	665982	R	YhhY-Amino Group Acetyl Transferase-(CT458)
CPn0576	667543	666494	R	prfB-Peptide Chain Release Factor 2 (natural UGA frame-shift)-(CT459)
CPn0576	667598	667530	R	prfB-(natural UGA frame-shift)
CPn0577	667895	668155	F	SWIB (YM74) complex protein-(CT460)
CPn0578	668406	669365	F	yaeI-phosphohydrolase-(CT461)
CPn0579	669361	669993	F	ygbP/yacM-Sugar Nucleotide Phosphorylase-(CT462)
CPn0580	669993	670793	F	truA-Pseudouridylate Synthase I-(CT463)
CPn0581	671434	670745	R	Phosphoglycolate Phosphatase-(CT464)
CPn0582	671503	672177	F	CT465 hypothetical protein
CPn0583	672400	672717	F	CT466 hypothetical protein
CPn0584	672707	673798	F	atoS/ntrB-2-Component Sensor-(CT467)
CPn0585	675817	673865	F	*similarity to Cps IncA_2
CPn0586	676026	677183	F	atoC/ntrC-2-Component Regulator-(CT468)
CPn0587	677441	678124	F	*yvyD_Bs conserved hypothetical protein
CPn0588	678084	678626	F	CT469 hypothetical protein
CPn0589	678640	679395	F	CT470 hypothetical protein
CPn0590	680112	679516	F	CT471 hypothetical protein
CPn0591	680373	681020	F	yagE family-(CT472)
CPn0592	681153	681461	F	yidD family-(CT473)
CPn0593	682476	681391	F	CT474 hypothetical protein
CPn0594	682583	684958	F	pheT-phenylalanyl tRNA Synthetase Beta-(CT475)
CPn0595	684958	685926	F	CT476 hypothetical protein
CPn0596	685939	686457	F	ada-methyltransferase-(CT477)
CPn0597	688215	686479	R	oppC_2-Oligopeptide Permease_2-(CT478)
CPn0598	689697	688219	R	oppB_2-Oligopeptide Permease_2-(CT479)
CPn0599	691802	689682	R	oppA_5-oligopeptide Binding Lipoprotein_5-(CT480)
CPn0600	692147	691827	R	
CPn0601	693053	692736	R	CT483 hypothetical protein
CPn0602	694105	693104	R	CT484 hypothetical protein
CPn0603	694205	695185	F	hemZ-Ferrochelatase-(CT485)
CPn0604	695945	695196	R	fliY-Glutamine Binding Protein-(CT486)
CPn0605	696707	696150	R	yhhF-Methylase -(CT487)
CPn0606	697444	696707	R	CT488 hypothetical protein
CPn0607	698895	697573	R	glgC-Glucose-1-P Adenyltransferase-(CT489)
CPn0608	699645	699016	R	*pyrF-Uridine 5'-Monophosphate Synthase (Ump Synthase)-truncated?
CPn0609	699705	699986	F	CT490 hypothetical protein
CPn0610	701420	700029	R	rho-Transcription Termination Factor-(CT491)
CPn0611	702025	701420	R	yacE-predicted phosphatase/kinase-(CT492)
CPn0612	704631	702022	R	polA-DNA Polymerase I-(CT493)
CPn0613	705656	704658	R	sohB-Protease-(CT494)
CPn0614	707402	705783	R	adt_2-ADP/ATP Translocase_2-(CT495)
CPn0615	708137	707634	R	pgsA_1-Glycerol-3-P Phosphatidyltransferase_1-(CT496)
CPn0616	708791	710137	F	dnaB-Replicative DNA Helicase-(CT497)
CPn0617	710484	712316	F	gidA-FAD-dependent oxidoreductase-(CT498)
CPn0618	712306	713010	F	lplA-Lipoate-Protein Ligase A-(CT499)
CPn0619	713444	713013	R	ndk-Nucleoside-2-P Kinase-(CT500)
CPn0620	714139	713519	R	ruvA-Holliday Junction Helicase-(CT501)
CPn0621	714647	714144	R	ruvC-Crossover Junction Endonuclease-(CT502)
CPn0622	715752	714793	R	CT503 hypothetical protein
CPn0623	716993	716163	R	CT504 hypothetical protein
CPn0624	718015	717011	R	gapA-Glyceraldehyde-3-P Dehydrogenase-(CT505)
CPn0625	718485	718060	R	rl17-L17 Ribosomal Protein-(CT506)
CPn0626	719616	718495	R	rpoA-RNA Polymerase Alpha-(CT507)
CPn0627	720038	719640	R	rs11-S11 Ribosomal Protein-(CT508)
CPn0628	720428	720063	R	rs13-S13 Ribosomal Protein-(CT509)
CPn0629	721857	720487	R	secY-Translocase-(CT510)
CPn0630	722316	721885	R	rl15-L15 Ribosomal Protein-(CT511)
CPn0631	722806	722312	R	rs5-S5 Ribosomal Protein-(CT512)
CPn0632	723195	722827	R	rl18-L18 Ribosomal Protein-(CT513)
CPn0633	723757	723209	R	rl6-L6 Ribosomal Protein-(CT514)
CPn0634	724185	723787	R	rs8-S8 Ribosomal Protein-(CT515)
CPn0635	724745	724206	R	rl5-L5 Ribosomal Protein-(CT516)
CPn0636	725082	724750	R	rl24-L24 Ribosomal Protein-(CT517)
CPn0637	725464	725099	R	rl14-L14 Ribosomal Protein-(CT518)
CPn0638	725747	725490	R	rs17-S17 Ribosomal Protein-(CT519)

CPn0639	725958	725743	R	rl29-L29 Ribosomal Protein-(CT520)
CPn0640	726377	725964	R	rl16-L16 Ribosomal Protein-(CT521)
CPn0641	727077	726409	R	rs3-S3 Ribosomal Protein-(CT522)
CPn0642	727428	727096	R	rl22-L22 Ribosomal Protein-(CT523)
CPn0643	727713	727450	R	rs19-S19 Ribosomal Protein-(CT524)
CPn0644	728573	727722	R	rl2-L2 Ribosomal Protein-(CT525)
CPn0645	728930	728598	R	rl23-L23 Ribosomal Protein-(CT526)
CPn0646	729621	728950	R	rl4-L4 Ribosomal Protein-(CT527)
CPn0647	730331	729657	R	rl3-L3 Ribosomal Protein-(CT528)
CPn0648	731603	730605	R	CT529 hypothetical protein
CPn0649	732672	731710	R	fmc-Methionyl tRNA Formyltransferase-(CT530)
CPn0650	733501	732665	R	lpxA-Acyl-Carrier UDP-GlcNAc -(CT531)
CPn0651	733975	733517	R	fabZ-Myristoyl-Acyl Carrier Dehydratase-(CT532)
CPn0652	734835	733990	R	lpxC-Myristoyl GlcNAc Deacetylase-(CT533)
CPn0653	736490	734868	R	cutE-Apolipoprotein N-Acetyltransferase-(CT534)
CPn0654	736967	736503	R	vdID/yciA-acyl-CoA Thioesterase-(CT535)
CPn0655	737847	737101	R	dnaQ_2-DNA Pol III Epsilon Chain_2-(CT536)
CPn0656	737872	738048	F	
CPn0657	738473	738051	R	yjeE (ATPase or Kinase)-(CT537)
CPn0658	739168	738455	R	CT538 hypothetical protein
CPn0659	739533	739838	F	trxA-Thioredoxin-(CT539)
CPn0660	740327	739860	R	spoU_2-rRNA Methylase_2-(CT540)
CPn0661	741100	740327	R	mip-FKBP-type peptidyl-prolyl cis-trans isomerase-(CT541)
CPn0662	742923	741172	R	aspS-Aspartyl tRNA Synthetase-(CT542)
CPn0663	744190	742901	R	hiss-Histidyl tRNA Synthetase-(CT543)
CPn0664	744757	744557	R	
CPn0665	745001	746365	F	uhpC-Hexosphosphate Transport -(CT544)
CPn0666	746388	750107	F	dnaE-DNA Pol III Alpha-(CT545)
CPn0667	751058	750177	R	predicted OMP [leader (17)]-(CT546)
CPn0668	751209	752162	F	CT547 hypothetical protein
CPn0669	752179	752775	F	CT548 hypothetical protein
CPn0670	752765	753196	F	rsbW-sigma regulatory factor-histidine kinase-(CT549)
CPn0671	753630	753205	R	CT550 hypothetical protein
CPn0672	753741	755048	F	dacF(pbp5)-D-Ala-D-Ala Carboxypeptidase-(CT551)
CPn0673	755287	755463	F	CT552 hypothetical protein
CPn0674	756668	755577	R	fmu-RNA Methyltransferase-(CT553)
CPn0675	757919	756768	R	CT696 hypothetical protein
CPn0676	759217	758051	R	homologous to CT695
CPn0677	760401	759256	R	
CPn0678	761320	760682	R	
CPn0679	762930	761725	R	pgk-Phosphoglycerate Kinase-(CT693)
CPn0680	764248	762971	R	ygo4-Phosphate Permease-(CT692)
CPn0681	764929	764258	R	CT691 hypothetical protein
CPn0682	764984	765955	F	dppD-ABC ATPase Dipeptide Transport-(CT690)
CPn0683	765948	766919	F	dppF-ABC ATPase Dipeptide Transport-(CT689)
CPn0684	768038	767181	R	spoJ/parB-Chromosome Partitioning Protein-(CT688)
CPn0685	768068	768217	F	
CPn0686	768361	768176	R	
CPn0687	768564	769214	F	CT482 hypothetical protein
CPn0688	769382	770137	F	CT481 hypothetical protein
CPn0689	771404	770187	R	yfhO_1-NifS-related Aminotransferase_1-(CT687)
CPn0690	772680	771436	R	ABC Transporter Membrane Protein-(CT686)
CPn0691	773452	772685	R	abcX-ABC Transporter ATPase-(CT685)
CPn0692	774912	773461	R	ABC Transporter-(CT684)
CPn0693	776256	775240	R	TPR Repeats (O-Linked GlcNAc Transferase similarity)-(CT683)
CPn0694	779599	776330	R	pbp2-PBP2-transglycolase/transpeptidase-(CT682)
CPn0695	780216	781382	F	ompA-Major Outer Membrane Protein-(CT681)
CPn0696	781769	782599	F	rs2-S2 Ribosomal Protein-(CT680)
CPn0697	782602	783447	F	tsf-Elongation Factor TS-(CT679)
CPn0698	783458	784201	F	pyrH-UMP Kinase-(CT679)
CPn0699	784182	784721	F	rrf-Ribosome Releasing Factor-(CT677)
CPn0700	785097	785609	F	CT676 hypothetical protein
CPn0701	785599	786672	F	karG-Arginine Kinase-(CT675)
CPn0702	789685	786929	R	yscC/gspD-Yop C/Gen Secretion Protein D-(CT674)
CPn0703	791190	789685	R	pknS-S/T Protein Kinase-(CT673)
CPn0704	792321	791209	R	flhN- Flagellar Motor Switch Domain/YscQ family-(CT672)
CPn0705	793173	792334	R	CT671 hypothetical protein
CPn0706	793683	793180	R	CT670 hypothetical protein
CPn0707	795029	793704	R	yscN-Yop N (Flagellar-Type ATPase)-(CT669)
CPn0708	795705	795034	R	CT668 hypothetical protein
CPn0709	796188	795742	R	CT667 hypothetical protein
CPn0710	796461	796210	R	CT666 hypothetical protein

CPn0711	796731	796486	R	CT665 hypothetical protein
CPn0712	799315	796781	R	FHA domain; homology to adenylate cyclase)-(CT664)
CPn0713	799721	799332	R	CT663 hypothetical protein
CPn0714	801107	800091	R	hemA-Glutamyl tRNA Reductase-(CT662)
CPn0715	801657	803462	F	gyrB_2-DNA Gyrase Subunit B_2-(CT661)
CPn0716	803469	804902	F	gyrA_2-DNA Gyrase Subunit A_2-(CT660)
CPn0717	805010	805306	F	CT656 hypothetical protein
CPn0718	805309	805626	F	CT657 hypothetical protein
CPn0719	805916	806890	F	sfhB-(Pseudouridine Synthase)-(CT658)
CPn0720	807003	807236	F	CT659 hypothetical protein
CPn0721	807683	808489	F	kdsA-KDO Synthetase-(CT655)
CPn0722	808489	808974	F	CT654 hypothetical protein
CPn0723	808984	809703	F	yhbG-ABC Transporter ATPase-(CT653)
CPn0724	810527	809706	R	
CPn0725	810811	810587	R	CT652.1 hypothetical protein
CPn0726	813372	810880	R	CT620 hypothetical protein
CPn0727	813577	816192	F	CT619 hypothetical protein
CPn0728	818477	816525	R	CHLPN 76kDa Homolog_1 (CT622)
CPn0729	819857	818592	R	CHLPN 76kDa Homolog_2 (CT623)
CPn0730	821603	819963	R	mviN-Integral Membrane Protein-(CT624)
CPn0731	821587	821760	F	
CPn0732	822098	822976	F	nfo-Endonuclease IV-(CT625)
CPn0733	823727	823101	R	rs4-S4 Ribosomal Protein-(CT626)
CPn0734	823944	824915	F	yceA-(CT627)
CPn0735	825668	825003	R	*pyrH/udk-Uridine Kinase (Uridine Monophosphokinase) (Pyrimidine Ribonucleoside Kinase).
CPn0736	827686	825992	R	ygeD-Efflux Protein-(CT641)
CPn0737	827685	830756	F	recC-Exodeoxyribonuclease V, Gamma-(CT640)
CPn0738	830746	833895	F	recB-Exodeoxyribonuclease V, Beta-(CT639)
CPn0739	834871	833861	R	CT638 hypothetical protein
CPn0740	836048	834864	R	tyrB-Aromatic AA Aminotransferase-(CT637)
CPn0741	838350	836185	R	greA-Transcription Elongation Factor-(CT636)
CPn0742	838463	838888	F	CT635 hypothetical protein
CPn0743	838962	840362	F	nqrA-Ubiquinone Oxidoreductase, Alpha-(CT634)
CPn0744	841384	840389	R	hemB-Porphobilinogen Synthase-(CT633)
CPn0745	841903	841742	R	
CPn0746	841975	843567	F	CT632 hypothetical protein
CPn0747	843675	843740	F	CT631 hypothetical protein
CPn0747	843725	843910	F	CT631 hypothetical protein (frame-shift)
CPn0748	844987	844121	R	ispA-Geranyl Transtransferase-(CT628)
CPn0749	845629	845006	R	glmU-UDP-GlcNAc Pyrophosphorylase-(CT629)
CPn0750	846411	845707	R	tctD/cpxR-HTH Transcriptional Regulatory Protein + Receiver Domain-(CT630)
CPn0751	846608	848434	F	CT651 hypothetical protein
CPn0752	848604	850082	F	recD_2-Exodeoxyribonuclease V, Alpha_2-(CT652)
CPn0753	851006	850161	R	
CPn0754	851336	851040	R	rs20-S20 Ribosomal Protein-(CT617)
CPn0755	851597	852799	F	CT616 hypothetical protein
CPn0756	852961	854676	F	rpoD-RNA Polymerase Sigma-66 -(CT615)
CPn0757	854733	855134	F	folX-Dihydroneopterin Aldolase-(CT614)
CPn0758	855110	856459	F	folP/dhpS-Dihydropteroate Synthase-(CT613)
CPn0759	856488	856997	F	folA-Dihydrofolate Reductase-(CT612)
CPn0760	856957	857694	F	CT611 hypothetical protein
CPn0761	857704	858375	F	CT610 hypothetical protein
CPn0762	859597	858539	R	recA-RecA recombination protein-(CT650)
CPn0763	860511	859972	R	ygfA-Formyltetrahydrofolate Cycloligase-(CT649)
CPn0764	861807	860524	R	CT648 hypothetical protein
CPn0765	862382	861801	R	CT647 hypothetical protein
CPn0766	863782	862394	R	CT646 hypothetical protein
CPn0767	863884	864177	F	CT645 hypothetical protein
CPn0768	864159	865163	F	yohI/nir3-predicted oxidoreductase -(CT644)
CPn0769	867733	865121	R	topA-DNA Topoisomerase I-Fused to SWI Domain-(CT643)
CPn0770	868340	869131	F	CT642 hypothetical protein
CPn0771	870463	869144	R	rpoN-RNA Polymerase Sigma-54-(CT609)
CPn0772	872385	870469	R	uvrD-DNA Helicase-(CT608)
CPn0773	872488	873195	F	ung-Uracil DNA Glycosylase-(CT607)
CPn0774	873195	873425	F	CT606.1 hypothetical protein
CPn0775	874031	873414	R	yggV family-(CT606)
CPn0776	874246	875487	F	CT605 hypothetical protein
CPn0777	875601	877178	F	groEL_2-heat shock protein-60 -(CT604)
CPn0778	877505	878092	F	tsa/ahpC-Thio-specific Antioxidant (TSA) Peroxidase-(CT603)
CPn0779	878481	878095	R	CT602 hypothetical protein

CPn0780	879205	878591	R	papQ/amiB-N-Acetylmuramoyl-L-Ala Amidase-(CT601)
CPn0781	879773	879198	R	pal-Peptidoglycan-Associated Lipoprotein-(CT600)
CPn0782	881065	879773	R	tolB-polysaccharide transporter-(CT599)
CPn0783	881885	881100	R	CT598 hypothetical protein
CPn0784	882296	881892	R	exbD-Biopolymer Transport Protein-(CT597)
CPn0785	882991	882296	R	exbB/tolQ-polysaccharide transporter-(CT596)
CPn0786	883185	885293	F	dsbD/xprA-Thio:disulfide Interchange Protein-(CT595)
CPn0787	885619	886401	F	yabD/ycfH-PHP superfamily (urease/pyrimidinase) hydrolase-(CT594)
CPn0788	886542	887432	F	sdhC-Succinate Dehydrogenase-(CT593)
CPn0789	887439	889316	F	sdhA-Succinate Dehydrogenase-(CT592)
CPn0790	889330	890103	F	sdhB-Succinate Dehydrogenase-(CT591)
CPn0791	893050	890111	R	CT590 hypothetical protein
CPn0792	894919	893108	R	CT589 hypothetical protein
CPn0793	896823	894919	R	rbsU-sigma regulatory family protein-PP2C phosphatase (RsbW antagonist)-(CT588)
CPn0794	897174	898004	F	
CPn0795	898128	899195	F	
CPn0796	899301	901340	F	
CPn0797	901600	902694	F	
CPn0798	902846	903856	F	
CPn0799	904986	903940	R	
CPn0800	906532	905249	R	eno-Enolase-(CT587)
CPn0801	908697	906727	R	uvrB-Exinuclease ABC Subunit B-(CT586)
CPn0802	909740	908709	R	trpS-Tryptophanyl tRNA Synthetase-(CT585)
CPn0803	910303	909752	R	CT584 hypothetical protein
CPn0804	911059	910310	R	gp6D-CHLTR Plasmid Paralog-(CT583)
CPn0805	911831	911067	R	minD-chromosome partitioning ATPase-CHLTR plasmid protein GP5D-(CT582)
CPn0806	913771	911867	R	thrS-Threonyl tRNA Synthetase-(CT581)
CPn0807	913971	914879	F	CT580 hypothetical protein
CPn0808	916287	914956	R	CT579 hypothetical protein
CPn0809	917785	916307	R	CT578 hypothetical protein
CPn0810	918184	917825	R	CT577 hypothetical protein
CPn0811	918900	918208	R	lcrH ₁ -Low Ca Response Protein H ₁ -(CT576)
CPn0812	919123	920862	F	mutL-DNA Mismatch Repair-(CT575)
CPn0813	920870	921934	F	pepP-Aminopeptidase P-(CT574)
CPn0814	922107	923357	F	CT573 hypothetical protein
CPn0815	923361	925622	F	gspD/pilQ-Gen. Secretion Protein D-(CT572)
CPn0816	925615	927102	F	gspE-Gen. Secretion Protein E-(CT571)
CPn0817	927115	928287	F	gspF-Gen. Secretion Protein F-(CT570)
CPn0818	928314	928682	F	predicted OMP [leader (16) peptide]-(CT569)
CPn0819	928689	929132	F	CT568 hypothetical protein
CPn0820	929120	929659	F	CT567 hypothetical protein
CPn0821	929667	930668	F	CT566 hypothetical protein
CPn0822	930756	931229	F	CT565 hypothetical protein
CPn0823	932367	931501	R	yscT/spaR-YopT Translocation T-(CT564)
CPn0824	932662	932378	R	yscS/fliQ-YopS/fliQ Translocation Protein-(CT563)
CPn0825	933594	932677	R	yscR-Yop Translocation R-(CT562)
CPn0826	934310	933612	R	yscL-Yop Translocation L-(CT561)
CPn0827	935264	934434	R	CT560 hypothetical protein
CPn0828	936271	935267	R	yscJ-Yop Translocation J-(CT559)
CPn0829	936744	937298	F	
CPn0830	937444	937959	F	
CPn0831	938267	938434	F	
CPn0832	939747	938827	R	lipA-Lipoate Synthetase-(CT558)
CPn0833	941129	939747	R	lpdA-Lipoamide Dehydrogenase-(CT557)
CPn0834	941553	942014	F	CT556 hypothetical protein
CPn0835	945689	942045	R	mot1 ₁ -SWI/SNF family helicase ₁ -(CT555)
CPn0836	946879	945722	R	brnQ-Amino Acid (Branched) Transport-(CT554)
CPn0837	947771	947145	R	nth-Enodnuclease III-(CT697)
CPn0838	949106	947781	R	thdF-Thiophene/Furan Oxidation Protein-(CT698)
CPn0839	949257	950159	F	psdD-Phosphatidylserine Decarboxylase-(CT699)
CPn0840	950222	951544	F	CT700 hypothetical protein
CPn0841	951731	954640	F	secA ₂ -Translocase SecA ₂ -(CT701)
CPn0842	954883	954710	R	CT702 hypothetical protein (frame-shift with 0843)
CPn0843	955191	954994	R	CT702 hypothetical protein
CPn0844	956730	955270	R	yphC-GTPase/GTP-binding protein-(CT703)
CPn0845	958079	956850	R	pcnB ₁ -Poly A Polymerase ₁ -(CT704)
CPn0846	959374	958112	R	clpX-CLP Protease ATPase-(CT705)
CPn0847	959995	959387	R	clpP-CLP Protease Subunit-(CT706)
CPn0848	961502	960177	R	tig/murI-Trigger Factor-peptidyl-prolyl isomerase-(CT707)
CPn0849	961788	965285	F	mot1 ₂ -SWI/SNF family helicase ₂ -(CT708)
CPn0850	965293	966390	F	mreB-Rod Shape Protein-Sugar Kinase-(CT709)

CPn0851	966396	968195	F	pckA-Phosphoenolpyruvate Carboxykinase-(CT710)
CPn0852	968316	970613	F	CT711 hypothetical protein
CPn0853	970637	971803	F	CT712 hypothetical protein
CPn0854	972837	971806	R	ompB-Outer Membrane Protein B-(CT713)
CPn0855	973995	972994	R	gpdA-Glycerol-3-P Dehydrogenase-(CT714)
CPn0856	975377	973995	R	AgX-1 Homolog-UDP-Glucose Pyrophosphorylase-(CT715)
CPn0857	975757	975392	R	CT716 hypothetical protein
CPn0858	977055	975757	R	fliI-Flagellum-specific ATP Synthase-(CT717)
CPn0859	977588	977055	R	CT718 hypothetical protein
CPn0860	978630	977608	R	fliF-Flagellar M-Ring Protein-(CT719)
CPn0861	979722	978925	R	nifU-NifU-related protein-(CT720)
CPn0862	980873	979722	R	yfhO_2-NifS-related protein_2-(CT721)
CPn0863	981514	980831	R	pgmA-Phosphoglycerate Mutase-(CT722)
CPn0864	981670	982374	F	yjbC-predicted pseudouridine synthase-(CT723)
CPn0865	982418	982942	F	CT724 hypothetical protein
CPn0866	983491	982916	R	birA-Biotin Synthetase-(CT725)
CPn0867	983423	984667	F	rodA-Rod Shape Protein-(CT726)
CPn0868	986643	984670	F	zntA/cadA-Metal Transport P-type ATPase-(CT727)
CPn0869	987401	986658	F	CT728 hypothetical protein
CPn0870	988728	987448	F	serS-Seryl tRNA Synthetase_2-(CT729)
CPn0871	988772	989899	F	ribD-Riboflavin Deaminase-(CT730)
CPn0872	989963	991216	F	ribA&ribB-GTP Cyclohydratase & DHBP Synthase -(CT731)
CPn0873	991233	991694	F	ribE-Ribityllumazine Synthase-(CT732)
CPn0874	993107	991749	F	CT733 hypothetical protein
CPn0875	993372	994022	F	CT734 hypothetical protein
CPn0876	994144	995517	F	dagA_2-D-Alanine/Glycine Permease_2-(CT735)
CPn0877	995533	995982	F	ybcL family-(CT736)
CPn0878	996654	995992	F	SET Domain protein-(CT737)
CPn0879	997439	996645	R	yycJ-metal dependent hydrolase-(CT738)
CPn0880	999861	997444	R	ftsK-Cell Division Protein FtsK-(CT739)
CPn0881	1005667	1006209	F	
CPn0882	1006268	1007404	F	
CPn0883	1008865	1007573	R	dmpP/nqr6-Phenolhydrolase/NADH ubiquinone oxidoreductase-(CT740)
CPn0884	1009359	1009009	R	CT741 hypothetical protein
CPn0885	1010635	1009433	R	ygcA-rRNA Methyltransferase-(CT742)
CPn0886	1011276	1010908	R	hctA-Histone-Like Developmental Protein-(CT743)
CPn0887	1011692	1014157	F	CHLTR possible phosphoprotein-(CT744)
CPn0888	1015423	1014119	R	hemG-protoporphyrinogen Oxidase-(CT745)
CPn0889	1016835	1015462	R	hemN_2-Coproporphyrinogen III Oxidase_2-(CT746)
CPn0890	1017805	1016819	R	hemE-Uroporphyrinogen Decarboxylase-(CT747)
CPn0891	1021073	1017819	R	mfd-Transcription-Repair Coupling-(CT748)
CPn0892	1023661	1021046	R	alaS-Alanyl tRNA Synthetase-(CT749)
CPn0893	1023894	1025888	F	tktB-Transketolase-(CT750)
CPn0894	1026766	1025888	R	amn-AMP Nucleosidase-(CT751)
CPn0895	1026988	1027557	F	efp_2-Elongation Factor P_2-(CT752)
CPn0896	1027595	1027822	F	CT753 hypothetical protein
CPn0897	1028737	1027853	R	(possible phosphohydrolase)-(CT754)
CPn0898	1030460	1028904	R	Mitochondrial HSP60 Chaperonin Homolog-(CT755)
CPn0899	1030875	1032215	F	murF-Muramoyl-DAP Ligase-(CT756)
CPn0900	1032235	1033281	F	mraY-Muramoyl-Pentapeptide Transferase-(CT757)
CPn0901	1033287	1034537	F	murd-Muramoylalanine-Glutamate Ligase-(CT758)
CPn0902	1034543	1035241	F	nlpD-Muramidase (invasin repeat family)-(CT759)
CPn0903	1035263	1036417	F	ftsW-Cell Division Protein FtsW-(CT760)
CPn0904	1036326	1037396	F	murG-Peptidoglycan Transferase-(CT761)
CPn0905	1037409	1039835	F	murC&ddlA-Muramate-Ala Ligase & D-Ala-D-Alam Ligase-(CT762)
CPn0906	1040340	1039915	R	CT763 hypothetical protein
CPn0907	1040780	1040445	R	*cutA Periplasmic Divalent Cation Tolerance Protein CutA (C-Type Cytochrome Biogenesis Protein)
CPn0908	1041589	1040780	R	CT764 hypothetical protein
CPn0909	1041637	1041966	F	rsbV_2-Sigma Factor Regulator_2-(CT765)
CPn0910	1041979	1043004	F	miaA-tRNA Pyrophosphate Transferase-(CT766)
CPn0911	1044043	1042985	R	Fe-S cluster oxidoreductase_2-(CT767)
CPn0912	1044129	1045760	F	CT768 hypothetical protein
CPn0913	1045760	1045945	F	
CPn0914	1045999	1046397	F	
CPn0915	1046461	1046817	F	ybeB-iojap superfamily ortholog-(CT769)
CPn0916	1046837	1048084	F	fabF-Acyl Carrier Protein Synthase-(CT770)
CPn0917	1048090	1048539	F	hydrolase/phosphatase homolog-(CT771)
CPn0918	1049223	1048579	R	ppa-Inorganic Pyrophosphatase-(CT772)
CPn0919	1049378	1050430	F	ldh-Leucine Dehydrogenase-(CT773)
CPn0920	1051405	1050431	R	cysQ-Sulfite Synthesis/biphosphate phosphatase-(CT774)
CPn0921	1051535	1052293	F	snGlycerol-3-P Acyltransferase-(CT775)

CPn0922	1052314	1053927	F	aas-Acylglycerophosphoethanolamine Acyltransferase-(CT776)
CPn0923	1053984	1055093	F	bioF_1-Oxononanoate Synthase_1-(CT777)
CPn0924	1057274	1055028	R	prpA-Primosomal Protein N' -(CT778)
CPn0925	1057900	1057226	R	CT779 hypothetical protein
CPn0926	1058060	1058557	F	Thioredoxin Disulfide Isomerase-(CT780)
CPn0927	1059809	1058670	R	*CHLPS 43 kDa protein homolog_2
CPn0928	1061008	1059884	R	*CHLPS 43 kDa protein homolog_3
CPn0929	1062292	1061186	R	*CHLPS 43 kDa protein homolog_4
CPn0930	1062857	1063330	F	
CPn0931	1064138	1065718	F	lysS-Lysyl tRNA Synthetase-(CT781)
CPn0932	1067142	1065721	R	cysS-CysteinyI tRNA Synthetase-(CT782)
CPn0933	1067535	1068578	F	predicted disulfide bond isomerase-(CT783)
CPn0934	1068942	1068526	R	rnpA-Ribonuclease P Protein Component-(CT784)
CPn0935	1069091	1068957	R	rl34-L34 Ribosomal Protein-(CT785)
CPn0936	1069336	1069470	F	rl36-L36 Ribosomal Protein-(CT786)
CPn0937	1069496	1069798	F	rl14-S14 Ribosomal Protein-(CT787)
CPn0938	1070322	1069849	R	CT788 hypothetical protein -[leader (60) peptide-periplasmic]
CPn0939	1070728	1071195	F	CT790 hypothetical protein
CPn0940	1073012	1071204	R	uvrC-Excinuclease ABC, Subunit C-(CT791)
CPn0941	1075501	1073018	R	mutS-DNA Mismatch Repair-(CT792)
CPn0942	1075985	1077754	F	dnaG/prim-DNA Primase-(CT794)
CPn0943	1077978	1078238	F	CT794.1 hypothetical protein
CPn0944	1078512	1078997	F	
CPn0945	1079070	1079660	F	CT795 hypothetical protein
CPn0946	1082786	1079745	R	glyQ-Glycyl tRNA Synthetase-(CT796)
CPn0947	1083442	1084059	F	pgsA_2-Glycerol-3-P-Phosphatidyltransferase_2-(CT797)
CPn0948	1085474	1084047	R	glgA-Glycogen Synthase-(CT798)
CPn0949	1085929	1086483	F	ctc-General Stress Protein-(CT799)
CPn0950	1086488	1087027	F	pth-Peptidyl tRNA Hydrolase-(CT800)
CPn0951	1087122	1087457	F	rs6-S6 Ribosomal Protein-(CT801)
CPn0952	1087478	1087723	F	rs18-S18 Ribosomal Protein-(CT802)
CPn0953	1087742	1088248	F	rl9-L9 Ribosomal Protein-(CT803)
CPn0954	1088286	1088708	F	ychB-Predicted Kinase-(CT804)
CPn0955	1088612	1089175	F	(frame-shift with 0954)
CPn0956	1089560	1090909	F	CT805 hypothetical protein
CPn0957	1093788	1090963	R	ide/ptr-Insulinase family/Protease III-(CT806)
CPn0958	1094785	1093793	R	plsB-Glycerol-3-P Acyltransferase-(CT807)
CPn0959	1096343	1094799	R	cafE-Axial Filament Protein-(CT808)
CPn0960	1096764	1097102	F	CT809 hypothetical protein
CPn0961	1097118	1097297	F	rl32-L32 Ribosomal Protein-(CT810)
CPn0962	1097316	1098275	F	plsX-FA/Phospholipid Synthesis Protein-(CT811)
CPn0963	1098398	1103224	F	pmp_21-Polymorphic Outer Membrane Protein D Family-(CT812)
CPn0964	1104758	1103301	R	
CPn0965	1106736	1104925	R	lpxB-Lipid A Disaccharide Synthase-(CT411)
CPn0966	1108037	1106748	R	pcnB_2-PolyA Polymerase_2-(CT410)
CPn0967	1108512	1109885	F	mrsA/pgm-Phosphoglucomutase-(CT815)
CPn0968	1109895	1111721	F	glmS-Glucosamine-Fructose-6-P Aminotransferase-(CT816)
CPn0969	1111812	1112999	F	0969-tyrP_1-Tyrosine Transport_1-(CT817) tyrP_1-Tyrosine Transport_1-(CT817)
CPn0970	1113461	1114648	F	0970-tyrP_2-Tyrosine Transport_2-(CT818) tyrP_2-Tyrosine Transport_2-(CT818)
CPn0971	1114702	1115415	F	yccA-Transport Permease-(CT819)
CPn0972	1116299	1115430	R	ftsY-Cell Division Protein FtsY-(CT820)
CPn0973	1116370	1117527	F	sucC-Succinyl-CoA Synthetase, Beta-(CT821)
CPn0974	1117544	1118422	F	sucD-Succinyl-CoA Synthetase, Alpha-(CT822)
CPn0975	1119104	1119637	F	
CPn0976	1120082	1121185	F	
CPn0977	1121371	1122402	F	
CPn0978	1122665	1123693	F	
CPn0979	1123980	1125443	F	htrA-DO Serine Protease-(CT823)
CPn0980	1126982	1125504	R	*similarity to Saccharomyces cerevisiae hypothetical 52.9KD protein
CPn0981	1127031	1129952	F	Zinc Metalloprotease (insulinase family)-(CT824)
CPn0982	1131194	1129962	R	yigN family-(CT825)
CPn0983	1132000	1131206	R	pssA-Glycerol-Serine Phosphatidyltransferase-(CT826)
CPn0984	1132379	1135510	F	nrdA-Ribonucleoside Reductase, Large Chain-(CT827)
CPn0985	1135534	1136571	F	nrdB-Ribonucleoside Reductase, Small Chain-(CT828)
CPn0986	1136724	1137395	F	yggH-predicted rRNA Methylase-(CT829)
CPn0987	1137516	1138115	F	ytgB-like predicted rRNA methylase-(CT830)
CPn0988	1138986	1138075	R	murB-UDP-N-Acetylenolpyruvoylglucosamine Reductase-(CT831)
CPn0989	1139495	1139016	R	CT832 hypothetical protein
CPn0990	1139883	1140440	F	infC-Initiation Factor 3-(CT833)
CPn0991	1140421	1140612	F	rl35-L35 Ribosomal Protein-(CT834)

CPn0992	1140634	1140996	F	rl20-L20 Ribosomal Protein-(CT835)
CPn0993	1141014	1142030	F	pheS-Phenylalanyl tRNA Synthetase, Alpha-(CT836)
CPn0994	1142398	1144440	F	CT837 hypothetical protein
CPn0995	1145512	1144415	R	CT838 hypothetical protein
CPn0996	1146589	1145519	R	CT839 hypothetical protein
CPn0997	1146708	1147664	F	mesJ-PP-loop superfamily ATPase-(CT840)
CPn0998	1147855	1150584	F	ftsH-ATP-dependent zinc protease-(CT841)
CPn0999	1152847	1150766	R	pnp-Polyribonucleotide Nucleotidyltransferase-(CT842)
CPn1000	1153157	1152891	R	rs15-S15 Ribosomal Protein-(CT843)
CPn1001	1153405	1153869	F	yfhC-cytosine deaminase-(CT844)
CPn1002	1153862	1154089	F	CT845 hypothetical protein
CPn1003	1154796	1154092	R	CT846 hypothetical protein
CPn1004	1155397	1154879	R	CT847 hypothetical protein
CPn1005	1155933	1155415	R	CT848 hypothetical protein
CPn1006	1156472	1155990	R	CT849 hypothetical protein
CPn1007	1156689	1156907	F	CT849.1 hypothetical protein
CPn1008	1156928	1158223	F	CT850 hypothetical protein
CPn1009	1159058	1158186	R	map-Methionine Aminopeptidase-(CT851)
CPn1010	1159672	1159067	R	CT852 hypothetical protein
CPn1011	1160306	1159902	R	CT853 hypothetical protein
CPn1012	1162193	1160421	R	yzcB-ABC transporter permease-(CT854)
CPn1013	1162245	1163624	F	fumC-Fumarate Hydratase-(CT855)
CPn1014	1165426	1163732	R	yhmH-Sulfate Transporter-(CT856)
CPn1015	1165634	1166893	F	CT857 hypothetical protein (possible IM protein)
CPn1016	1167042	1168898	F	CT858 hypothetical protein
CPn1017	1169006	1169935	F	lytB-Metalloprotease-(CT859)
CPn1018	1169898	1170629	F	
CPn1019	1172128	1170638	R	CT860 hypothetical protein
CPn1020	1173679	1172150	R	CT861 hypothetical protein
CPn1021	1174213	1173698	R	lcrH_2-Low Calcium Response_2-(CT862)
CPn1022	1175673	1174216	R	CT863 hypothetical protein
CPn1023	1176035	1176331	F	
CPn1024	1177236	1176334	R	xerD-Integrase/recombinase-(CT864)
CPn1025	1177302	1178879	F	pgi-Glucose-6-P Isomerase-(CT378)
CPn1026	1178997	1179137	F	ltuA-(CT377)
CPn1027	1179175	1180755	F	
CPn1028	1181016	1181999	F	mdhC-Malate Dehydrogenase-(CT376)
CPn1029	1182008	1182844	F	
CPn1030	1183886	1182843	R	predicted D-amino acid dehydrogenase-(CT375)
CPn1031	1185552	1184098	R	arcD-Arginine/Ornithine Antiporter-(CT374)
CPn1032	1186150	1185566	R	CT373 hypothetical protein
CPn1033	1187500	1186187	R	CT372 hypothetical protein
CPn1034	1188517	1187732	R	Predicted OMP_1 (CT371) [leader (18) peptide]
CPn1035	1190000	1188570	R	AroE-Shikimate 5-Dehydrogenase-(CT370)
CPn1036	1191135	1189984	R	AroB-Dehydroquinase Synthase-(CT369)
CPn1037	1192199	1191123	R	AroC-Chorismate Synthase-(CT368)
CPn1038	1192726	1192199	R	aroL-Shikimate Kinase II-(CT367)
CPn1039	1193999	1192665	R	aroA-Phosphoshikimate Vinyltransferase-(CT366)
CPn1040	1194741	1194073	R	
CPn1041	1195994	1194726	R	*bioA-Adenosylmethionine-8-Amino-7-Oxononanoate Aminotransferase
CPn1042	1196590	1195934	R	*bioD-dethiobiotin synthetase
CPn1043	1197717	1196572	R	bioF_2-Oxononanoate Synthase_2
CPn1044	1198691	1197699	R	*bioB-Biotin Synthase
CPn1045	1199590	1198901	R	*conserved hypothetical bacterial membrane protein
CPn1046	1200675	1199590	R	*Tryptophan Hydroxylase
CPn1047	1200552	1201343	F	dapB-Dihydrodipicolinate Reductase-(CT364)
CPn1048	1201606	1202604	F	asd-Aspartate Dehydrogenase-(CT363)
CPn1049	1202595	1203914	F	lysC-Aspartokinase III-(CT362)
CPn1050	1203926	1204798	F	dapA-Dihydrodipicolinate Synthase-(CT361)
CPn1051	1204962	1205270	F	
CPn1052	1205417	1206169	F	
CPn1053	1206153	1206701	F	
CPn1054	1207034	1209466	F	
CPn1055	1209694	1210521	F	
CPn1056	1210527	1211228	F	
CPn1057	1211497	1213596	F	CT356 hypothetical protein
CPn1058	1213748	1214836	F	CT355 hypothetical protein
CPn1059	1214848	1215678	F	kgsA-Dimethyladenosine Transferase-(CT354)
CPn1060	1217658	1215727	R	dxs/tkt-Transketolase-(CT331)
CPn1061	1217920	1217666	R	CT330 hypothetical protein
CPn1062	1219820	1218159	R	xseA-Exodeoxyribonuclease VII-(CT329)
CPn1063	1219951	1220712	F	tpiS-Triosephosphate Isomerase-(CT328)

CPn1064	1220719	1220895	F	
CPn1065	1221095	1220928	R	
CPn1066	1221135	1221488	F	
CPn1067	1221735	1222292	F	def-Polypeptide Deformylase-(CT353)
CPn1068	1223258	1222365	R	rnhB_2-Ribonuclease HII_2-(CT008)
CPn1069	1223513	1223941	F	yfgA-HTH Transcriptional Regulator-(CT009)
CPn1070	1225511	1224144	R	
CPn1071	1227324	1225885	R	
CPn1072	1227969	1228835	F	
CPn1073	1229011	1229832	F	Predicted OMP_2 -(CT371)

Table 2 (Supplemental Data) Functional Assignments of *C. pneumoniae* Coding Sequences *C. trachomatis* genes are shown in parentheses

5	Amino Acid Biosynthesis			
	<i>Aromatic Family</i>			
	1039	(CT366)	aroA	Phosphoshikimate Vinyltransferase
	1036	(CT369)	aroB	Dehydroquinate Synthase
	1037	(CT368)	aroC	Chorismate Synthase
10	1035	(CT370)	aroE	Shikimate 5-Dehydrogenase
	0484	(CT382)	aroG	Deoxyheptonate Aldolase
	1038	(CT367)	aroL	Shikimate Kinase II
	0740	(CT637)	tyrB	Aromatic AA Aminotransferase
	<i>Aspartate Family (lysine)</i>			
15	1048	(CT363)	asd	Aspartate Dehydrogenase
	1050	(CT361)	dapA	Dihydropicolinate Synthase
	1047	(CT364)	dapB	Dihydropicolinate Reductase
	0519	(CT430)	dapF	Dianinopimelate Epimerase
	1049	(CT362)	lysC	Aspartokinase III
20	<i>Serine Family</i>			
	0433	(CT282)	gcsH	Glycine Cleavage System H Protein
	0521	(CT432)	glyA	Serine Hydroxymethyltransferase
	<i>Base & Nucleotide Metabolism</i>			
25	0171		guaA	GMP Synthase
	0172		guaB	Inosine 5'-Monophosphate Dehydrogenase
	0608			Undine 5'-Monophosphate Synthase
	0735			Undine Kinase
	0244	(CT128)	adk	Adenylyate Kinase
	0894	(CT751)	amn	AMP Nucleosidase
30	0568	(CT452)	cmk	CMP Kinase
	0392	(CT039)	dcd	dCTP Deaminase
	0059	(CT292)	dut	dUTP Nucleotidohydrolase
	0120	(CT030)	gmK	GMP Kinase
	0619	(CT500)	ndk	Nucleoside-2-P Kinase
35	0984	(CT827)	nrdA	Ribonucleoside Reductase, Large Chain
	0985	(CT828)	nrdB	Ribonucleoside Reductase, Small Chain
	0236	(CT183)	pyrG	CTP Synthetase
	0698	(CT678)	pyrH	UMP Kinase
40	0273	(CT188)	tdk	Thymidylate Kinase
	0659	(CT539)	trxA	Thioredoxin
	0314	(CT099)	trxB	Thioredoxin Reductase
	1001	(CT844)	yfhC	Cytosine Deaminase
	Biosynthesis of Cofactors			
45	<i>Biotin, Lipoate & Ubiquinone</i>			
	1041		bioA	Adenosylmethionine-8-Amino-7-Oxononanoate Aminotransferase
	1044		bioB	Biotin Synthase
	1042		bioD	Dethiobiotin Synthetase
	0923	(CT777)	bioF_1	Oxononanoate Synthase_1
50	1043	(CT777)	bioF_2	Oxononanoate Synthase_2
	0866	(CT725)	birA	Biotin Synthetase
	0748	(CT628)	ispA	Geranyl Transtransferase
	0832	(CT558)	lipA	Lipoate Synthetase

	0265	(CT219)	ubiA	Benzoate Octaphenyltransferase
	0264	(CT220)	ubiD	Phenylacrylate Decarboxylase
	0515	(CT428)	ubiE	Ubiquinone Methyltransferase
	<i>Folic Acid</i>			
5	0759	(CT612)	folA	Dihydrofolate Reductase
	0335	(CT078)	folD	Methylene Tetrahydrofolate Dehydrogenase
	0758	(CT613)	folP	Dihydropteroate Synthase
	0757	(CT614)	folX	Dihydroneopterin Aldolase
	0763	(CT649)	ygfA	Formyltetrahydrofolate Cycloligase
10	<i>Porphyrin</i>			
	0714	(CT662)	hemA	Glutamyl tRNA Reductase
	0744	(CT633)	hemB	Porphobilinogen Synthase
	0052	(CT299)	hemC	Porphobilinogen Deaminase
	0890	(CT747)	hemE	Uroporphyrinogen Decarboxylase
15	0888	(CT745)	hemG	protoporphyrinogen Oxidase
	0138	(CT210)	hemL	Glutamate-1-Semialdehyde-2,1-Aminomutase
	0380	(CT052)	hemN_1	Coproporphyrinogen III Oxidase_1
	0889	(CT746)	hemN_2	Coproporphyrinogen III Oxidase_2
	0603	(CT485)	hemZ	Ferrochetalase
20	<i>Riboflavin</i>			
	0872	(CT731)	nbA&nbB	GTP Cyclohydrtase & DHBP Synthase
	0532	(CT405)	nbC	Riboflavin Synthase
	0871	(CT730)	nbD	Riboflavin Deaminase
	0873	(CT732)	nbE	Ribityllumazine Synthase
25	0320	(CT093)	nbF	FAD Synthase
	<i>Cell Envelope</i>			
	<i>Fatty Acid & Phospholipid Metabolism</i>			
	0161	(CT206)		(predicted acyltransferase family)
30	0922	(CT776)	aas	Acylglycerophosphoethanolamine Acyltransferase
	0414	(CT265)	accA	AcCoA Carboxylase/Transferase Alpha
	0183	(CT123)	accB	Biotin Carboxyl Carrier Protein
	0182	(CT124)	accC	Biotin Carboxylase
	0058	(CT293)	accD	AcCoA Carboxylase/Transferase Beta
35	0295	(CT236)	acpP	Acyl Carrier Protein
	0313	(CT100)	acpS	Acyl-carrier Protein Synthase
	0567	(CT451)	cdsA	Phosphatidate Cytidyltransferase
	0297	(CT238)	fabD	Malonyl Acyl Carrier Transcyclase
	0916	(CT770)	fabF	Acyl Carrier Protein Synthase
40	0296	(CT237)	fabG	Oxoacyl (Carrier Protein) Reductase
	0298	(CT239)	fabH	Oxoacyl Carrier Protein Synthase III
	0406	(CT104)	fabI	Enoyl-Acyl-Carrier Protein Reductase
	0651	(CT532)	fabZ	Myristoyl-Acyl Carrier Dehydratase
	0098	(CT010)	htrB	Acyltransferase
45	0271	(CT136)		Lysophospholipase Esterase
	0615	(CT496)	pgsA_1	Glycerol-3-P Phosphatidyltransferase_1
	0947	(CT797)	pgsA_2	Glycerol-3-P Phosphatidyltransferase_2
	0958	(CT807)	plsB	Glycerol-3-P Acyltransferase
	0569	(CT453)	plsC	Glycerol-3-P Acyltransferase
50	0962	(CT811)	plsX	FA/Phospholipid Synthesis Protein
	0839	(CT699)	psdD	Phosphatidylserine Decarboxylase
	0983	(CT826)	pssA	Glycerol-Serine Phosphatidyltransferase
	0921	(CT775)		snGlycerol-3-P Acyltransferase
	0654	(CT535)	yciA	Acyl-CoA Thioesterase
55	0877	(CT736)	ybcL	CT736 Hypothetical Protein

LPS

0154	(CT208)	<i>gseA</i>	KDO Transferase
0721	(CT655)	<i>kdsA</i>	KDO Synthetase
0235	(CT182)	<i>kdsB</i>	Deoxyoctulosonic Acid Synthetase
0650	(CT531)	<i>lpxA</i>	Acyl-Carrier UDP-GlcNAc O-Acyltransferase
0965	(CT411)	<i>lpxB</i>	Lipid A Disaccharide Synthase
0652	(CT533)	<i>lpxC</i>	Mynstoyl GlcNAc Deacetylase
0302	(CT243)	<i>lpxD</i>	UDP Glucosamine N-Acyltransferase

Membrane Proteins, Lipoproteins & Porins

0310	(CT251)	<i>60IM</i>	60kDa Inner Membrane Protein
0556	(CT442)	<i>crpA</i>	15kDa Cysteine-Rich Protein
0653	(CT534)	<i>cutE</i>	Apolipoprotein N-Acetyltransferase
0311	(CT252)	<i>lgt</i>	Prolipoprotein Diacylglycerol Transferase
0558	(CT444)	<i>omcA</i>	9kDa-Cysteine-Rich Lipoprotein
0557	(CT443)	<i>omcB</i>	60kDa Cysteine-Rich OMP
0695	(CT681)	<i>ompA</i>	Major Outer Membrane Protein
0854	(CT713)	<i>ompB</i>	Outer Membrane Protein B
0781	(CT600)	<i>pal</i>	Peptidoglycan-Associated Lipoprotein
0300	(CT241)	<i>yaeT</i>	Omp85 Homolog

Peptidoglycan

0417	(CT268)	<i>amiA</i>	N-Acetylmuramoyl Alanine Amidase
0780	(CT601)	<i>amiB</i>	N-Acetylmuramoyl-L-Ala Amidase
0672	(CT551)	<i>dacF</i>	D-Ala-D-Ala Carboxypeptidase
0968	(CT816)	<i>glmS</i>	Glucosamine-Fructose-6-P Aminotransferase
0749	(CT629)	<i>glmU</i>	UDP-GlcNAc Pyrophosphorylase
0900	(CT757)	<i>mraY</i>	Muramoyl-Pentapeptide Transferase
0571	(CT455)	<i>murA</i>	UDP-N-Acetylglucosamine Transferase
0988	(CT831)	<i>murB</i>	UDP-N-Acetylenolpyruvoylglucosamine Reductase
0905	(CT762)	<i>murC&ddlA</i>	Muramate-Ala Ligase & D-Ala-D-Ala Ligase
0901	(CT758)	<i>murD</i>	Muramoylalanine-Glutamate Ligase
0418	(CT269)	<i>murE</i>	N-Acetylmuramoylalanylglutamyl DAP Ligase
0899	(CT756)	<i>murF</i>	Muramoyl-DAP Ligase
0904	(CT761)	<i>murG</i>	Peptidoglycan Transferase
0902	(CT759)	<i>nlpD</i>	Muramidase (invasin repeat family)
0694	(CT682)	<i>pbp2</i>	PBP2-Transglycolase/Transpeptidase
0419	(CT270)	<i>pbp3</i>	Transglycolase/Transpeptidase
0421	(CT272)	<i>yabC</i>	PBP2B Family Methyltransferase

Cellular Processes

Cell Division

0959	(CT808)	<i>cafE</i>	Axial Filament Protein
0880	(CT739)	<i>ftsK</i>	Cell Division Protein FtsK
0903	(CT760)	<i>ftsW</i>	Cell Division Protein FtsW
0972	(CT820)	<i>ftsY</i>	Cell Division Protein FtsY
0617	(CT498)	<i>gidA</i>	FAD-dependent Oxidoreductase
0805	(CT582)	<i>minD</i>	Chromosome Partitioning ATPase
0850	(CT709)	<i>mreB</i>	Rod Shape Protein-Sugar Kinase
0867	(CT726)	<i>rodA</i>	Rod Shape Protein
0684	(CT688)	<i>parB</i>	Chromosome Partitioning Protein

Detoxification

0057	(CT294)	<i>sodM</i>	Superoxide Dismutase (Mn)
0778	(CT603)	<i>ahpC</i>	Thio-specific Antioxidant (TSA) Peroxidase

Signal Transduction

0148	(CT145)		S/T Protein Kinase
0584	(CT467)	<i>atoS</i>	Two-Component Sensor
0294	(CT235)		cAMP-Dependent Protein Kinase Regulatory Subunit
0712	(CT664)		(FHA domain)

5	0599	(CT480)	oppA_5	Oligopeptide Binding Lipoprotein_5
	0199	(CT199)	oppB_1	Oligopeptide Permease_1
	0598	(CT479)	oppB_2	Oligopeptide Permease_2
	0200	(CT200)	oppC_1	Oligopeptide Permease_1
	0597	(CT478)	oppC_2	Oligopeptide Permease_2
10	0201	(CT201)	oppD	Oligopeptide Transport ATPase
	0202	(CT202)	oppF	Oligopeptide Transport ATPase
	0231	(CT180)	tauB	ABC Transport ATPase (Nitrate/Fe)
	0782	(CT599)	tolB	Macromolecule Transporter
	0969	(CT817)	tyrP_1	Tyrosine Transport_1
15	0970	(CT818)	tyrP_2	Tyrosine Transport_2
	0665	(CT544)	uhpC	Hexosphosphate Transport
	0282	(CT216)	xasA	Amino Acid Transporter
	0207	(CT204)	ybhI	dicarboxylate Translocator
	0971	(CT819)	yccA	Transport Permease
20	0248	(CT152)	ycfV	ABC Transporter ATPase
	1014	(CT856)	ychM	Sulfate Transporter
	0736	(CT641)	ygeD	Efflux Protein
	0680	(CT692)	ygo4	Phosphate Permease
	0723	(CT653)	yhbG	ABC Transporter ATPase
25	0023	(CT348)	yjjK	ABC Transporter Protein ATPase
	0127	(CT034)	ytfF	Cationic Amino Acid Transporter
	0349	(CT067)	ytgA	Solute Protein Binding Family
	0348	(CT068)	ytgB	ABC transporter ATPase
	0347	(CT069)	ytgC	Integral Membrane Protein
30	0346	(CT070)	ytgD	Integral Membrane Protein
	1012	(CT854)	yzeB	ABC Transporter Permease
	0868	(CT727)	zntA	Metal Transport P-type ATPase
	0279			Possible ABC Transporter Permease Protein
	0543	(CT417)		(Metal Transport Protein)
35	0692	(CT684)		ABC Transporter
	0542	(CT416)		ABC Transporter ATPase
	0690	(CT686)		ABC Transporter Membrane Protein
	0541	(CT415)		solute binding protein
				<i>Type-III Secretion</i>
40	0323	(CT090)	lcrD	Low Calcium Response D
	0324	(CT089)	lcrE	Low Calcium Response E
	0811	(CT576)	lcrH_1	Low Ca Response Protein H_1
	1021	(CT862)	lcrH_2	Low Calcium Response_2
	0325	(CT088)	sycE	Secretion Chaperone
45	0702	(CT674)	yscC	Yop C/Gen Secretion Protein D
	0828	(CT559)	yscJ	Yop Translocation J
	0826	(CT561)	yscL	Yop Translocation L
	0707	(CT669)	yscN	Yop N (Flagellar-Type ATPase)
	0825	(CT562)	yscR	Yop Translocation R
50	0824	(CT563)	yscS	YopS Translocation Protein
	0823	(CT564)	yscT	YopT Translocation T
	0322	(CT091)	yscU	Yop Translocation Protein U
				Central Intermediary Metabolism
				<i>Glycogen Metabolism</i>
55	0856	(CT715)		UDP-Glucose Pyrophosphorylase
	0948	(CT798)	glgA	Glycogen Synthase
	0475	(CT866)	glgB	Glucan Branching Enzyme
	0607	(CT489)	glgC	Glucose-1-P Adenylyltransferase
	0307	(CT248)	glgP	Glycogen Phosphorylase
	0388	(CT042)	glgX	Glycogen Hydrolase (debranching)

0326	(CT087)	malQ	Glucanotransferase
0851	(CT710)	pckA	Phosphoenolpyruvate Carboxykinase

Phosphorous & Sulfur

0548	(CT435)	cysJ	Sulfite Reductase
0920	(CT774)	cysQ	Sulfite Synthesis/Biphosphate Phosphatase
0025	(CT346)	atsA	Sulphohydrolase
0918	(CT772)	ppa	Inorganic Pyrophosphatase

DNA Replication, Modification, Repair & Recombination

DNA Mismatch Repair

0505			3-Methyladenine DNA Glycosylase
0812	(CT575)	mutL	DNA Mismatch Repair
0941	(CT792)	mutS	DNA Mismatch Repair
0402	(CT107)	mutY	Adenine Glycosylase
0732	(CT625)	nfo	Endonuclease IV
0837	(CT697)	nth	Endonuclease III

DNA Modification

0596	(CT477)	ada	Methyltransferase
0114	(CT024)	hemK	A/G-specific Methylase
0891	(CT748)	mfd	Transcription-Repair Coupling
0620	(CT501)	ruvA	Holliday Junction Helicase
0390	(CT040)	ruvB	Holliday Junction Helicase
0621	(CT502)	ruvC	Crossover Junction Endonuclease
0053	(CT298)	sms	Sms Protein
0773	(CT607)	ung	Uracil DNA Glycosylase
1062	(CT329)	xseA	Exodeoxynuclease VII

DNA Recombination

0762	(CT650)	recA	RecA Recombination Protein
0738	(CT639)	recB	Exodeoxynuclease V, Beta
0737	(CT640)	recC	Exodeoxynuclease V, Gamma
0123	(CT033)	recD_1	Exodeoxynuclease V (Alpha Subunit)_1
0752	(CT652)	recD_2	Exodeoxynuclease V, Alpha_2
0339	(CT074)	recF	ABC Superfamily ATPase
0340	(CT074)		(frame-shift with 0339)
0563	(CT447)	recJ	ssDNA Exonuclease
0299	(CT240)	recR	Recombination Protein

DNA Replication

0309	(CT250)	dnaA_1	Replication Initiation Protein_1
0424	(CT275)	dnaA_2	Replication Initiation Factor_2
0616	(CT497)	dnaB	Replicative DNA Helicase
0666	(CT545)	dnaE	DNA Pol III Alpha
0942	(CT794)	dnaG	DNA Primase
0338	(CT075)	dnaN	DNA Pol III (Beta)
0410	(CT261)	dnaQ_1	DNA Pol III Epsilon Chain_1
0655	(CT536)	dnaQ_2	DNA Pol III Epsilon Chain_2
0040	(CT334)	dnaX_1	DNA Pol III Gamma and Tau_1
0272	(CT187)	dnaX_2	DNA Pol III Gamma and Tau_2
0149	(CT146)	dnlJ	DNA Ligase
0274	(CT189)	gyrA_1	DNA Gyrase Subunit A_1
0716	(CT660)	gyrA_2	DNA Gyrase Subunit A_2
0275	(CT190)	gyrB_1	DNA Gyrase Subunit B_1
0715	(CT661)	gyrB_2	DNA Gyrase Subunit B_2
0416	(CT267)	himD	Integration Host Factor Alpha
0612	(CT493)	polA	DNA Polymerase I
0924	(CT778)	priA	Primosomal Protein N'
0386	(CT044)	ssb	SS DNA Binding Protein

Genes of the *Salmonella enterica* serovar *Agona* genome

5	0835	(CT555)		SWI/SNF family helicase_1
	0849	(CT708)		SWI/SNF family helicase_2
	0769	(CT643)	topA	DNA Topoisomerase I-Fused to SWI Domain
	0024	(CT347)	xerC	Integrase/recombinase
	1024	(CT864)	xerD	Integrase/recombinase

Eukaryotic-Type Chromatin Factors

	0886	(CT743)	hctA	Histone-Like Developmental Protein
	0384	(CT046)	hctB	Histone-like Protein 2
10	0878	(CT737)		SET Domain protein
	0577	(CT460)		SWIB (YM74) Complex Protein

UVR Exinuclease Repair System

	0096	(CT333)	uvrA	Excinuclease ABC Subunit A
	0801	(CT586)	uvrB	Exinuclease ABC Subunit B
	0940	(CT791)	uvrC	Excinuclease ABC, Subunit C
15	0772	(CT608)	uvrD	DNA Helicase

Energy Metabolism

Aerobic

20	0855	(CT714)	gpdA	Glycerol-3-P Dehydrogenase
	0743	(CT634)	nqrA	Ubiquinone Oxidoreductase, Alpha
	0427	(CT278)	nqr2	NADH (Ubiquinone) Dehydrogenase
	0428	(CT279)	nqr3	NADH (Ubiquinone) Oxidoreductase, Gamma
	0429	(CT280)	nqr4	NADH (Ubiquinone) Reductase 4
	0430	(CT281)	nqr5	NADH (Ubiquinone) Reductase 5
25	0883	(CT740)	nqr6	Phenolhydrolase/NADH (Ubiquinone) Oxidoreductase 6

ATP Biogenesis and metabolism

	0351	(CT065)	adt_1	ADP/ATP Translocase_1
	0614	(CT495)	adt_2	ADP/ATP Translocase_2
30	0088	(CT308)	atpA	ATP Synthase Subunit A
	0089	(CT307)	atpB	ATP Synthase Subunit B
	0090	(CT306)	atpD	ATP Synthase Subunit D
	0086	(CT310)	atpE	ATP Synthase Subunit E
	0091	(CT305)	atpI	ATP Synthase Subunit I
	0092	(CT304)	atpK	ATP Synthase Subunit K
35	0860	(CT719)	flaF	Flagellar M-Ring Protein

Electron Transport Chain

	0102	(CT013)	cydA	Cytochrome Oxidase Subunit I
	0103	(CT014)	cydB	Cytochrome Oxidase Subunit II
40	0364	(CT059)		Ferredoxin
	0084	(CT312)		Predicted Ferredoxin

Glycolysis & Gluconeogenesis

	0281	(CT215)	dhnA	Predicted 1,6-Fructose Biphosphate Aldolase
	0800	(CT587)	eno	Enolase
45	0624	(CT505)	gapA	Glyceraldehyde-3-P Dehydrogenase
	0056	(CT295)	mrsA	Phosphomannomutase
	0967	(CT815)	pgm	Phosphoglucomutase
	0160	(CT207)	pfkA_1	Fructose-6-P Phosphotransferase_1
	0208	(CT205)	pfkA_2	Fructose-6-P Phosphotransferase_2
50	1025	(CT378)	pgi	Glucose-6-P Isomerase
	0679	(CT693)	pgk	Phosphoglycerate Kinase
	0863	(CT722)	pgmA	Phosphoglycerate Mutase
	0097	(CT332)	pyk	Pyruvate Kinase
	1063	(CT328)	tpiS	Tnophosphate Isomerase

Pentose Phosphate Pathway

55	0239	(CT186)	devB	Glucose-6-P Dehydrogenase (DevB family)
	1060	(CT331)	dxs	Transketolase

	0360	(CT063)	gnd	6-Phosphogluconate Dehydrogenase
	0185	(CT121)	rpe	Ribulose-P Epimerase
	0141	(CT213)	rpiA	Ribose-5-P Isomerase A
5	0083	(CT313)	tal	Transaldolase
	0893	(CT750)	tktB	Transketolase
	0238	(CT185)	zwf	Glucose-6-P Dehydrogenase
	<i>Pyruvate Dehydrogenase</i>			
	0833	(CT557)	lpdA	Lipoamide Dehydrogenase
10	0436	(CT285)	lplA_1	Lipoate Protein Ligase-Like Protein
	0618	(CT499)	lplA_2	Lipoate-Protein Ligase A
	0033	(CT340)	pdhA&B	Oxoisovalerate Dehydrogenase α/β Fusion
	0304	(CT245)	pdhA	Pyruvate Dehydrogenase Alpha
	0305	(CT246)	pdhB	Pyruvate Dehydrogenase Beta
15	0306	(CT247)	pdhC	Dihydrolipoamide Acetyltransferase
	<i>TCA Cycle</i>			
	0495	(CT390)	aspC	Aspartate Aminotransferase
	1013	(CT855)	fumC	Fumarate Hydratase
	1028	(CT376)	mdhC	Malate Dehydrogenase
20	0789	(CT592)	sdhA	Succinate Dehydrogenase
	0790	(CT591)	sdhB	Succinate Dehydrogenase
	0788	(CT593)	sdhC	Succinate Dehydrogenase
	0378	(CT054)	sucA	Oxoglutarate Dehydrogenase
	0377	(CT055)	sucB_1	Dihydrolipoamide Succinyltransferase_1
25	0527	(CT400)	sucB_2	Dihydrolipoamide Succinyltransferase_2
	0973	(CT821)	sucC	Succinyl-CoA Synthetase, Beta
	0974	(CT822)	sucD	Succinyl-CoA Synthetase, Alpha
	Protein Folding, Assembly & Modification			
	<i>Chaperones</i>			
30	0949	(CT799)	ctc	General Stress Protein
	0534	(CT407)	dksA	DnaK Suppressor
	0032	(CT341)	dnaJ	Heat Shock Protein J
	0503	(CT396)	dnaK	Hsp-70
	0134	(CT110)	groEL_1	Hsp-60_1
35	0777	(CT604)	groEL_2	Hsp-60_2
	0898	(CT755)	groEL_3	Hsp-60_3
	0135	(CT111)	groES	10KDa Chaperonin
	0502	(CT395)	grpE	HSP-70 Cofactor
40	0661	(CT541)	mip	FKBP-type Peptidyl-prolyl Cis-Trans Isomerase
	<i>Proteases</i>			
	0144	(CT113)	clpB	Clp Protease ATPase
	0437	(CT286)	clpC	ClpC Protease
	0520	(CT431)	clpP_1	CLP Protease
45	0847	(CT706)	clpP_2	CLP Protease Subunit
	0846	(CT705)	clpX	CLP Protease ATPase
	0269	(CT138)		Dipeptidase
	0998	(CT841)	ftsH	ATP-dependent Zinc Protease
	0030	(CT343)	gcp_1	O-Sialoglycoprotein Endopeptidase_1
50	0194	(CT197)	gcp_2	O-Sialoglycoprotein Endopeptidase_2
	0979	(CT823)	htrA	DO Serine Protease
	0957	(CT806)	ide	Insulinase family/Protease III
	0027	(CT344)	lon	Lon ATP-dependent Protease
	1017	(CT859)	lytB	Metalloprotease
55	1009	(CT851)	map	Methionine Aminopeptidase
	0385	(CT045)	pepA	Leucyl Aminopeptidase A
	0136	(CT112)	pepF	Oligopeptidase

0813	(CT574)	pepP	Aminopeptidase P
0613	(CT494)	sohB	Protease
0555	(CT441)	tsp	Tail-Specific Protease
0344	(CT072)	yaeL	Metalloprotease
0981	(CT824)		Zinc Metalloprotease (insulinase family)

Protein Isomerases

0227	(CT176)	dsbB	Disulfide bond Oxidoreductase
0786	(CT595)	dsbD	Thio disulfide Interchange Protein
0228	(CT177)	dsbG	Disulfide Bond Chaperone
0933	(CT783)		Predicted Disulfide Bond Isomerase
0926	(CT780)		Thioredoxin Disulfide Isomerase

Transcription

RNA Degradation

0999	(CT842)	pnp	Polynucleotide Nucleotidyltransferase
0054	(CT297)	mrc	Ribonuclease III
0119	(CT029)	mhb_1	Ribonuclease HII_1
1068	(CT008)	mhb_2	Ribonuclease HII_2
0934	(CT784)	mpA	Ribonuclease P Protein Component
0504	(CT397)	vacB	Ribonuclease Family

RNA Elongation & Termination Factors

0741	(CT636)	greA	Transcription Elongation Factor
0316	(CT097)	nusA	N Utilization Protein A
0076	(CT320)	nusG	Transcriptional Antitermination
0845	(CT704)	pcnB_1	Poly A Polymerase_1
0966	(CT410)	pcnB_2	PolyA Polymerase_2
0610	(CT491)	rho	Transcription Termination Factor

RNA Methylases

0674	(CT553)	fmu	RNA Methyltransferase
1059	(CT354)	kgsA	Dimethyladenosine Transferase
0187	(CT133)		Predicted Methylase
0530	(CT403)	spoU_1	rRNA Methylase_1
0660	(CT540)	spoU_2	rRNA Methylase_2
0117	(CT027)	trmD	tRNA (Guanine N-1)-Methyltransferase
0885	(CT742)	ygcA	rRNA Methyltransferase
0986	(CT829)	yggH	Predicted rRNA Methylase
0987	(CT830)	ytgB	Predicted rRNA Methylase

RNA Modification

0649	(CT530)	fnt	Methionyl tRNA Formyltransferase
0910	(CT766)	miaA	tRNA Pyrophosphate Transferase
0719	(CT658)	sfhB	Predicted Pseudouridine Synthase
0219	(CT193)	tgt	Queuine tRNA Ribosyl Transferase
0580	(CT463)	truA	Pseudouridylylase Synthase I
0319	(CT094)	truB	tRNA Pseudouridine Synthase
0403	(CT106)	yceC	Predicted Pseudouridine Synthetase Family
0864	(CT723)	yjbc	Predicted Pseudouridine Synthase

RNA Polymerase & Transcription Regulators

0586	(CT468)	atoC	Two-Component Regulator
0362	(CT061)	rpsD	Sigma-28/WhiG Family
0501	(CT394)	hrcA	HTH Transcriptional Repressor
0793	(CT588)	rsbU	Sigma Regulatory Family Protein—PP2C Phosphatase (RsbW Antagonist)
0626	(CT507)	rpoA	RNA Polymerase Alpha
0081	(CT315)	rpoB	RNA Polymerase Beta
0082	(CT314)	rpoC	RNA Polymerase Beta'
0756	(CT615)	rpoD	RNA Polymerase Sigma-66
0771	(CT609)	rpoN	RNA Polymerase Sigma-54
0511	(CT424)	rsbV_1	Sigma Regulatory Factor_1
0909	(CT765)	rsbV_2	Sigma Factor Regulator_2
0670	(CT549)	rsbW	Sigma Regulatory Factor-Histidine Kinase
0750	(CT630)	tctD	HTH Transcriptional Regulatory Protein + Receiver Domain
1069	(CT009)	yfgA	HTH Transcriptional Regulator

Translation

Amino Acyl tRNA Synthesis

0892	(CT749)	alaS	Alanyl tRNA Synthetase
0570	(CT454)	argS	Arginyl tRNA Transferase
0662	(CT542)	aspS	Aspartyl tRNA Synthetase

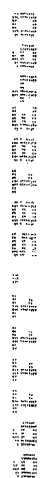
	0932	(CT782)	cysS	Cysteinyl tRNA Synthetase
	0003	(CT003)	gatA	Glu tRNA Gln Amidotransferase (A subunit)
	0004	(CT004)	gatB	Glu tRNA Gln Amidotransferase (B Subunit)
	0002	(CT002)	gatC	Glu tRNA Gln Amidotransferase (C subunit)
5	0560	(CT445)	glxX	Glutamyl-tRNA Synthetase
	0946	(CT796)	glyQ	Glycyl tRNA Synthetase
	0663	(CT543)	hisS	Histidyl tRNA Synthetase
	0109	(CT019)	ileS	Isoleucyl-tRNA Synthetase
	0153	(CT209)	leuS	Leucyl tRNA Synthetase
10	0931	(CT781)	lysS	Lysyl tRNA Synthetase
	0122	(CT032)	metG	Methionyl-tRNA Synthetase
	0993	(CT836)	pheS	Phenylalanyl tRNA Synthetase, Alpha
	0594	(CT475)	pheT	Phenylalanyl tRNA Synthetase Beta
	0500	(CT393)	proS	Prolyl tRNA Synthetase
15	0870	(CT729)	serS	Seryl tRNA Synthetase_2
	0806	(CT581)	thrS	Threonyl tRNA Synthetase
	0802	(CT585)	trpS	Tryptophanyl tRNA Synthetase
	0361	(CT062)	tyrS	Tyrosyl tRNA Synthetase
	0094	(CT302)	valS	Valyl tRNA Synthetase
20	<i>Peptide Chain Initiation, Elongation & Termination</i>			
	1067	(CT353)	def	Polypeptide Deformylase
	0184	(CT122)	efp_1	Elongation Factor P_1
	0895	(CT752)	efp_2	Elongation Factor P_2
	0550	(CT437)	fusA	Elongation Factor G
25	0073	(CT323)	infA	Initiation Factor IF-1
	0317	(CT096)	infB	Initiation Factor-2
	0990	(CT833)	infC	Initiation Factor 3
	0113	(CT023)	pfrA	Peptide Chain Releasing Factor 1
	0576	(CT459)	prfB	Peptide Chain Release Factor 2
30	0950	(CT800)	pth	Peptidyl tRNA Hydrolase
	0318	(CT095)	rbfA	Ribosome Binding Factor A
	0699	(CT677)	rrf	Ribosome Releasing Factor
	0697	(CT679)	tsf	Elongation Factor TS
	0074	(CT322)	tufA	Elongation Factor Tu
35	<i>Ribosomal Proteins</i>			
	0078	(CT318)	rl1	L1 Ribosomal Protein
	0644	(CT525)	rl2	L2 Ribosomal Protein
	0647	(CT528)	rl3	L3 Ribosomal Protein
	0646	(CT527)	rl4	L4 Ribosomal Protein
40	0635	(CT516)	rl5	L5 Ribosomal Protein
	0633	(CT514)	rl6	L6 Ribosomal Protein
	0080	(CT316)	rl7	L7/L12 Ribosomal Protein
	0953	(CT803)	rl9	L9 Ribosomal Protein
	0079	(CT317)	rl10	L10 Ribosomal Protein
45	0077	(CT319)	rl11	L11 Ribosomal Protein
	0247	(CT125)	rl13	L13 Ribosomal Protein
	0637	(CT518)	rl14	L14 Ribosomal Protein
	0630	(CT511)	rl15	L15 Ribosomal Protein
	0640	(CT521)	rl16	L16 Ribosomal Protein
50	0625	(CT506)	rl17	L17 Ribosomal Protein
	0632	(CT513)	rl18	L18 Ribosomal Protein
	0118	(CT028)	rl19	L19 Ribosomal Protein
	0992	(CT835)	rl20	L20 Ribosomal Protein
	0546	(CT420)	rl21	L21 Ribosomal Protein
55	0642	(CT523)	rl22	L22 Ribosomal Protein
	0645	(CT526)	rl23	L23 Ribosomal Protein

5	0636	(CT517)	rl24	L24 Ribosomal Protein
	0545	(CT419)	rl27	L27 ribosomal protein
	0327	(CT086)	rl28	L28 Ribosomal Protein
	0639	(CT520)	rl29	L29 Ribosomal Protein
	0112	(CT022)	rl31	L31 Ribosomal Protein
	0961	(CT810)	rl32	L32 Ribosomal Protein
	0250	(CT150)	rl33	L33 Ribosomal Protein
	0935	(CT785)	rl34	L34 Ribosomal Protein
10	0991	(CT834)	rl35	L35 Ribosomal Protein
	0936	(CT786)	rl36	L36 Ribosomal Protein
	0315	(CT098)	rs1	S1 Ribosomal Protein
	0696	(CT680)	rs2	S2 Ribosomal Protein
	0641	(CT522)	rs3	S3 Ribosomal Protein
15	0733	(CT626)	rs4	S4 Ribosomal Protein
	0631	(CT512)	rs5	S5 Ribosomal Protein
	0951	(CT801)	rs6	S6 Ribosomal Protein
	0551	(CT438)	rs7	S7 Ribosomal Protein
	0634	(CT515)	rs8	S8 Ribosomal Protein
20	0246	(CT126)	rs9	S9 Ribosomal Protein
	0549	(CT436)	rs10	S10 Ribosomal Protein
	0627	(CT508)	rs11	S11 Ribosomal Protein
	0552	(CT439)	rs12	S12 Ribosomal Protein
	0628	(CT509)	rs13	S13 Ribosomal Protein
25	0937	(CT787)	rs14	S14 Ribosomal Protein
	1000	(CT843)	rs15	S15 Ribosomal Protein
	0116	(CT026)	rs16	S16 Ribosomal Protein
	0638	(CT519)	rs17	S17 Ribosomal Protein
	0952	(CT802)	rs18	S18 Ribosomal Protein
30	0643	(CT524)	rs19	S19 Ribosomal Protein
	0754	(CT617)	rs20	S20 Ribosomal Protein
	0031	(CT342)	rs21	S21 Ribosomal Protein

35 Other Categories

<i>Chlamydia-Specific Proteins</i>				
	0561	(CT446)	Euo	CHLPS Euo Protein
	0804	(CT583)	Gp6D	CHLTR Plasmid Paralog
	0186	(CT119)		Similarity to IncA_1
40	0291	(CT232)	incB	Inclusion Membrane Protein B
	0292	(CT233)	incC	Inclusion Membrane Protein C
	1026	(CT377)		LtuA Protein
	0333	(CT080)		LtuB Protein
45	0005	(CT871)	pmp_1	Polymorphic Outer Membrane Protein G Family
	0013	(CT871)	pmp_2	Polymorphic Outer Membrane Protein G Family
	0014	(CT871)	pmp_3	Polymorphic Outer Membrane Protein G Family
	0015	(CT871)	pmp_3	PMP_3 (frame-shift with 0014)
	0016	(CT874)	pmp_4	Polymorphic Outer Membrane Protein G Family
	0017	(CT871)	pmp_4	PMP_4 (frame-shift with 0016)
50	0018	(CT874)	pmp_5	Polymorphic Outer Membrane Protein G Family
	0019	(CT871)	pmp_5	PMP_5 (frame-shift with 0018)
	0444	(CT871)	pmp_6	Polymorphic Outer Membrane Protein G/I Family
	0445	(CT871)	pmp_7	Polymorphic Outer Membrane Protein G Family
	0446	(CT871)	pmp_8	Polymorphic Outer Membrane Protein G Family
55	0447	(CT871)	pmp_9	Polymorphic Outer Membrane Protein G/I Family
	0450	(CT871)	pmp_10	Polymorphic Outer Membrane Protein G Family
	0449	(CT871)	pmp_10	PMP_10 (Frame-shift with 0450)

5	0451	(CT871)	pmp_11	Polymorphic Outer Membrane Protein G Family
	0452	(CT874)	pmp_12	Polymorphic Outer Membrane Protein (truncated) A/I Family
	0453	(CT871)	pmp_13	Polymorphic Outer Membrane Protein G Family
	0454	(CT872)	pmp_14	Polymorphic Outer Membrane Protein H Family
	0466	(CT869)	pmp_15	Polymorphic Outer Membrane Protein E Family
10	0467	(CT869)	pmp_16	Polymorphic Outer Membrane Protein E Family
	0468	(CT869)	pmp_17	Polymorphic Outer Membrane Protein E Family
	0469	(CT869)	pmp_17	PMP_17 (Frame-shift with 0468)
	0470	(CT869)	pmp_17	PMP_17 (Frame-shift with 0469)
	0471	(CT870)	pmp_18	Polymorphic Outer Membrane Protein E/F Family
15	0539	(CT412)	pmp_19	Polymorphic Membrane Protein A Family
	0540	(CT413)	pmp_20	Polymorphic Membrane Protein B Family
	0963	(CT812)	pmp_21	Polymorphic Membrane Protein D Family
	0562			CHLPS 43 kDa Protein Homolog_1
	0927			CHLPS 43 kDa Protein Homolog_2
20	0928			CHLPS 43 kDa Protein Homolog_3
	0929			CHLPS 43 kDa Protein Homolog_4
	0728	(CT622)		CHLPS 76kDa Homolog_1 (CT622)
	0729	(CT623)		CHLPS 76kDa Homolog_2 (CT623)
	0133	(CT109)		CHLPS Hypothetical Protein
	0332	(CT081)		CHLPS T2 Protein
<i>Miscellaneous Enzymes/Conserved Proteins</i>				
25	0193		argR	Possible Arginine Repressor
	1046			Aromatic Amino Acid Hydroxylase
	0232			Similarity to 5'-Methylthioadenosine Nucleosidase
	0128	(CT035)		Biotin Protein Ligase
	0513	(CT426)		Fe-S Oxidoreductase_1
30	0911	(CT767)		Fe-S Oxidoreductase_2
	0373	(CT057)	gcpE	GcpE Protein
	0407	(CT103)		HAD Superfamily Hydrolase/Phosphatase
	0917	(CT771)		Hydrolase/Phosphatase Homolog
	0488	(CT385)	ycfF	HIT Family Hydrolase
35	0701	(CT675)	karG	Arginine Kinase
	0526	(CT399)	kpsF	GutQ/KpsF Family Sugar-P Isomerase
	0919	(CT773)	ldh	Leucine Dehydrogenase
	0022	(CT349)	maf	Maf protein
	0997	(CT840)	mesJ	PP-loop superfamily ATPase
40	0151	(CT148)	mhpA	Monooxygenase
	0730	(CT624)	mviN	Integral Membrane Protein
	0861	(CT720)		NifU-Related Protein
	0479	(CT380)	phnP	Metal Dependent Hydrolase
	0106	(CT015)	phoH	ATPase
45	0329	(CT084)		Phospholipase D Superfamily
	0435	(CT284)		Phospholipase D Superfamily
	0581	(CT464)		Phosphoglycolate Phosphatase
	0897	(CT754)		Predicted Phosphohydrolase
	0509	(CT422)		Predicted Metalloenzyme
50	1030	(CT375)		Predicted D-Amino Acid Dehydrogenase
	0531	(CT404)		SAM Dependent Methyltransferase
	0337	(CT076)	smpB	Small Protein B
	0394	(CT256)	tlyC_1	CBS Domain Protein (Hemolysin Homolog)_1
	0510	(CT423)	tlyC_2	CBS Domains (Hemolysin Homolog)_2
55	0382	(CT048)	yabC	SAM-Dependent Methyltransferase
	0787	(CT594)	yabD	PHP Superfamily (Urease/Pyrimidinase) Hydrolase
	0611	(CT492)	yacE	Predicted Phosphatase/Kinase
	0579	(CT462)	yacM	Sugar Nucleotide Phosphorylase
	0578	(CT461)	yaeI	Phosphohydrolase



5	0345	(CT071)	yaeM	CT071 Hypothetical Protein
	0566	(CT450)	yaeS	YaeS family Hypothetical Protein
	0591	(CT472)	yagE	YagE family
	0039	(CT335)	ybaB	YbaB family Hypothetical Protein
	0101	(CT012)	ybbP	YbbP family Hypothetical Protein
10	0915	(CT769)	ybeB	iojap Superfamily Ortholog
	0137	(CT108)	ybgI	ACR family
	0529	(CT402)	ycaH	ATPase
	0438	(CT287)	ycbF	PP-loop Superfamily ATPase
	0734	(CT627)	yceA	YceA Hypothetical Protein
15	0954	(CT804)	ychB	Predicted Kinase
	0261	(CT217)	ydaO	PP-Loop Superfamily ATPase
	0245	(CT127)	ydhO	Polysaccharide Hydrolase-Invasin Repeat Family
	0573	(CT457)	yebC	YebC Family Hypothetical Protein
	0689	(CT687)	yfhO_1	NifS-related Aminotransferase_1
20	0862	(CT721)	yfhO_2	NifS-related Aminotransferase_2
	0547	(CT434)	ygbB	YgbB Family Hypothetical Protein
	0237	(CT184)	yggF	YggF Family Hypothetical Protein
	0775	(CT606)	yggV	YggV Family Hypothetical Protein
	0396	(CT258)	yhfO_3	NifS-related Aminotransferase_3
25	0605	(CT487)	yhhF	Predicted Methylase
	0575	(CT458)	yhhY	Amino Group Acetyl Transferase
	0592	(CT473)	yidD	YidD Family
	0982	(CT825)	yigN	YigN Family Hypothetical Protein
	0657	(CT537)	yjeE	YjeE Hypothetical Protein
30	0768	(CT644)	yohI	YohI Predicted Oxidoreductase
	0336	(CT077)	yojL	YojL Hypothetical Protein
	0217	(CT140)	ypdP	YpdP Hypothetical Protein
	0140	(CT212)	yqdE	YqdE Hypothetical Protein
	0263	(CT221)	yqfU	YqfU Hypothetical Protein
35	0139	(CT211)	yqgE	YqgE Hypothetical Protein
	0270	(CT137)	ywiC	SuA5 Superfamily-related Protein
	0879	(CT738)	yyeJ	Metal Dependent Hydrolase

Homologs to CHLTR Hypothetical Coding Genes

40	0001	(CT001)	CT001 Hypothetical Protein
	0020	(CT351)	CT351 Hypothetical Protein
	0021	(CT350)	CT350 Hypothetical Protein
	0026	(CT345)	CT345 Hypothetical Protein
	0035	(CT339)	CT339 Hypothetical Protein
45	0036	(CT338)	CT338 Hypothetical Protein
	0055	(CT296)	CT296 Hypothetical Protein
	0062	(CT289)	CT289 Hypothetical Protein
	0065	(CT288)	CT288 Hypothetical Protein
	0068	(CT360)	CT360 Hypothetical Protein
50	0071	(CT325)	CT325 Hypothetical Protein
	0072	(CT324)	CT324 Hypothetical Protein
	0085	(CT311)	CT311 Hypothetical Protein
	0087	(CT309)	CT309 Hypothetical Protein
	0093	(CT303)	CT303 Hypothetical Protein
55	0100	(CT011)	CT011 Hypothetical Protein
	0104	(CT017)	CT017 Hypothetical Protein
	0105	(CT016)	CT016 Hypothetical Protein
	0107	(CT058)	CT058 Hypothetical Protein_1
	0108	(CT018)	CT018 Similarity
	0111	(CT021)	CT021 Hypothetical Protein
	0121	(CT031)	CT031 Hypothetical Protein

	0129	(CT036)	CT036 Similarity
	0145	(CT114)	CT114 Hypothetical Protein
	0150	(CT147)	CT147 Hypothetical Protein
5	0152	(CT149)	CT149 Hypothetical Protein
	0176	(CT153)	CT153 Hypothetical Protein
	0188	(CT132)	CT132 Hypothetical Protein
	0189	(CT131)	CT131 Hypothetical Protein
	0206	(CT203)	CT203 Hypothetical Protein
10	0229	(CT178)	CT178 Hypothetical Protein
	0230	(CT179)	CT179 Hypothetical Protein
	0234	(CT181)	CT181 Hypothetical Protein
	0249	(CT151)	CT151 Hypothetical Protein
	0253	(CT144)	CT144 Hypothetical Protein_1
	0254	(CT143)	CT143 Hypothetical Protein_1
15	0255	(CT142)	CT142 Hypothetical Protein_1
	0256	(CT144)	CT144 Hypothetical Protein_2
	0257	(CT143)	CT143 Hypothetical Protein_2
	0259	(CT142)	CT142 Hypothetical Protein_2
20	0276	(CT191)	CT191 Hypothetical Protein
	0288	(CT195)	CT195 Hypothetical Protein
	0293	(CT234)	CT234 Hypothetical Protein
	0301	(CT242)	CT368 Hypothetical Protein
	0303	(CT244)	CT244 Hypothetical Protein
	0308	(CT249)	CT249 Similarity
25	0312	(CT101)	CT101 Hypothetical Protein
	0328	(CT085)	CT085 Hypothetical Protein
	0330	(CT083)	CT083 Hypothetical Protein
	0331	(CT082)	CT082 Hypothetical Protein
	0334	(CT079)	CT079 Similarity
30	0342	(CT073)	CT073 Hypothetical Protein
	0343	(CT073)	(frame-shift with 0342?)
	0350	(CT066)	CT066 Hypothetical Protein
	0369	(CT058)	CT058 Hypothetical Protein_2
	0370	(CT058)	CT058 Hypothetical Protein_3
35	0374	(CT056)	CT056 Hypothetical Protein
	0379	(CT053)	CT053 Hypothetical Protein
	0381	(CT326)	CT326 Similarity
	0383	(CT047)	CT047 Hypothetical Protein
	0387	(CT043)	CT043 Hypothetical Protein
40	0389	(CT041)	CT041 Hypothetical Protein
	0393	(CT038)	CT038 Hypothetical Protein
	0395	(CT257)	CT257 Hypothetical Protein
	0399	(CT253)	CT253 Hypothetical Protein
	0400	(CT254)	CT254 Hypothetical Protein
45	0401	(CT255)	CT255 Hypothetical Protein
	0405	(CT105)	CT105 Hypothetical Protein
	0408	(CT102)	CT102 Hypothetical Protein
	0409	(CT260)	CT260 Hypothetical Protein
	0411	(CT262)	CT262 Hypothetical Protein
50	0412	(CT263)	CT263 Hypothetical Protein
	0415	(CT266)	CT266 Hypothetical Protein
	0420	(CT271)	CT271 Hypothetical Protein
	0422	(CT273)	CT273 Hypothetical Protein
	0423	(CT274)	CT274 Hypothetical Protein
55	0425	(CT276)	CT276 Hypothetical Proteins
	0426	(CT277)	CT277 Similarity
	0434	(CT283)	CT283 Hypothetical Protein

	0441	(CT007)	CT007 Hypothetical Protein
	0442	(CT006)	CT006 Hypothetical Protein
	0443	(CT005)	CT005 Hypothetical Protein
5	0474	(CT365)	CT365 Hypothetical Protein
	0476	(CT865)	CT865 Hypothetical Protein
	0480	(CT383)	CT383 Hypothetical Protein
	0485	(CT382)	CT382.1 Hypothetical Protein
	0487	(CT384)	CT384 Hypothetical Protein
10	0489	(CT386)	CT386 Hypothetical Protein
	0490	(CT387)	CT387 Hypothetical Protein
	0491	(CT389)	CT389 Hypothetical Protein
	0496	(CT391)	CT391 Hypothetical Protein
	0497	(CT388)	CT388 Hypothetical Protein
15	0506	(CT421)	CT421 Hypothetical Protein
	0507	(CT421)	CT421 1 Hypothetical Protein
	0508	(CT421)	CT421 2 Hypothetical Protein
	0512	(CT425)	CT425 Hypothetical Protein
	0514	(CT427)	CT427 Hypothetical Protein
20	0518	(CT429)	CT429 Hypothetical Protein
	0522	(CT433)	CT433 Hypothetical Protein
	0525	(CT398)	CT398 Hypothetical Protein
	0533	(CT406)	CT406 Hypothetical Protein
	0537	(CT814)	CT814 1 Hypothetical Protein
25	0538	(CT814)	CT814 Hypothetical Protein
	0554	(CT440)	CT440 Hypothetical Protein
	0559	(CT441)	CT441.1 Hypothetical Protein
	0565	(CT449)	CT449 Hypothetical Protein
	0572	(CT456)	CT456 Hypothetical Protein
30	0582	(CT465)	CT465 Hypothetical Protein
	0583	(CT466)	CT466 Hypothetical Protein
	0588	(CT469)	CT469 Hypothetical Protein
	0589	(CT470)	CT470 Hypothetical Protein
	0590	(CT471)	CT471 Hypothetical Protein
35	0593	(CT474)	CT474 Hypothetical Protein
	0595	(CT476)	CT476 Hypothetical Protein
	0601	(CT483)	CT483 Hypothetical Protein
	0602	(CT484)	CT484 Hypothetical Protein
	0606	(CT488)	CT488 Hypothetical Protein
40	0609	(CT490)	CT490 Hypothetical Protein
	0622	(CT503)	CT503 Hypothetical Protein
	0623	(CT504)	CT504 Hypothetical Protein
	0648	(CT529)	CT529 Hypothetical Protein
	0658	(CT538)	CT538 Hypothetical Protein
45	0667	(CT546)	CT546 Hypothetical Protein
	0668	(CT547)	CT547 Hypothetical Protein
	0669	(CT548)	CT548 Hypothetical Protein
	0671	(CT550)	CT550 Hypothetical Protein
	0673	(CT552)	CT552 Hypothetical Protein
50	0675	(CT696)	CT696 Hypothetical Protein
	0676	(CT695)	CT695 Similarity
	0681	(CT691)	CT691 Hypothetical Protein
	0687	(CT482)	CT482 Hypothetical Protein
	0688	(CT481)	CT481 Hypothetical Protein
55	0700	(CT676)	CT676 Hypothetical Protein
	0705	(CT671)	CT671 Hypothetical Protein
	0706	(CT670)	CT670 Hypothetical Protein
	0708	(CT668)	CT668 Hypothetical Protein

	0709	(CT667)	CT667 Hypothetical Protein
	0710	(CT666)	CT666 Hypothetical Protein
	0711	(CT665)	CT665 Hypothetical Protein
5	0713	(CT663)	CT663 Hypothetical Protein
	0717	(CT656)	CT656 Hypothetical Protein
	0718	(CT657)	CT657 Hypothetical Protein
	0720	(CT659)	CT659 Hypothetical Protein
	0722	(CT654)	CT654 Hypothetical Protein
10	0725	(CT652)	CT652.1 Hypothetical Protein
	0726	(CT620)	CT620 Hypothetical Protein
	0727	(CT619)	CT619 Hypothetical Protein
	0739	(CT638)	CT638 Hypothetical Protein
	0742	(CT635)	CT635 Hypothetical Protein
15	0746	(CT632)	CT632 Hypothetical Protein
	0747	(CT631)	CT631 Hypothetical Protein
	0751	(CT651)	CT651 Hypothetical Protein
	0755	(CT616)	CT616 Hypothetical Protein
	0760	(CT611)	CT611 Hypothetical Protein
20	0761	(CT610)	CT610 Hypothetical Protein
	0764	(CT648)	CT648 Hypothetical Protein
	0765	(CT647)	CT647 Hypothetical Protein
	0766	(CT646)	CT646 Hypothetical Protein
	0767	(CT645)	CT645 Hypothetical Protein
25	0770	(CT642)	CT642 Hypothetical Protein
	0774	(CT606)	CT606.1 Hypothetical Protein
	0776	(CT605)	CT605 Hypothetical Protein
	0779	(CT602)	CT602 Hypothetical Protein
	0783	(CT598)	CT598 Hypothetical Protein
30	0791	(CT590)	CT590 Hypothetical Protein
	0792	(CT589)	CT589 Hypothetical Protein
	0803	(CT584)	CT584 Hypothetical Protein
	0807	(CT580)	CT580 Hypothetical Protein
	0808	(CT579)	CT579 Hypothetical Protein
35	0809	(CT578)	CT578 Hypothetical Protein
	0810	(CT577)	CT577 Hypothetical Protein
	0814	(CT573)	CT573 Hypothetical Protein
	0818	(CT569)	CT569 Hypothetical Protein
	0819	(CT568)	CT568 Hypothetical Protein
40	0820	(CT567)	CT567 Hypothetical Protein
	0821	(CT566)	CT566 Hypothetical Protein
	0822	(CT565)	CT565 Hypothetical Protein
	0827	(CT560)	CT560 Hypothetical Protein
	0834	(CT556)	CT556 Hypothetical Protein
45	0840	(CT700)	CT700 Hypothetical Protein
	0842	(CT702)	CT702 Hypothetical Protein
	0843	(CT702)	CT702 Hypothetical Protein
	0852	(CT711)	CT711 Hypothetical Protein
	0853	(CT712)	CT712 Hypothetical Protein
50	0857	(CT716)	CT716 Hypothetical Protein
	0859	(CT718)	CT718 Hypothetical Protein
	0865	(CT724)	CT724 Hypothetical Protein
	0869	(CT728)	CT728 Hypothetical Protein
	0874	(CT733)	CT733 Hypothetical Protein
	0875	(CT734)	CT734 Hypothetical Protein
55	0884	(CT741)	CT741 Hypothetical Protein
	0887	(CT744)	CHLTR Possible Phosphoprotein
	0896	(CT753)	CT753 Hypothetical Protein

	0906	(CT763)	CT763 Hypothetical Protein
	0908	(CT764)	CT764 Hypothetical Protein
	0912	(CT768)	CT768 Hypothetical Protein
5	0925	(CT779)	CT779 Hypothetical Protein
	0938	(CT788)	CT788 Hypothetical Protein
	0939	(CT790)	CT790 Hypothetical Protein
	0943	(CT794)	CT794.1 Hypothetical Protein
	0945	(CT795)	CT795 Hypothetical Protein
10	0956	(CT805)	CT805 Hypothetical Protein
	0960	(CT809)	CT809 Hypothetical Protein
	0989	(CT832)	CT832 Hypothetical Protein
	0994	(CT837)	CT837 Hypothetical Protein
	0995	(CT838)	CT838 Hypothetical Protein
	0996	(CT839)	CT839 Hypothetical Protein
15	1002	(CT845)	CT845 Hypothetical Protein
	1003	(CT846)	CT846 Hypothetical Protein
	1004	(CT847)	CT847 Hypothetical Protein
	1005	(CT848)	CT848 Hypothetical Protein
	1006	(CT849)	CT849 Hypothetical Protein
20	1007	(CT849)	CT849 1 Hypothetical Protein
	1008	(CT850)	CT850 Hypothetical Protein
	1010	(CT852)	CT852 Hypothetical Protein
	1011	(CT853)	CT853 Hypothetical Protein
	1015	(CT857)	CT857 Hypothetical Protein
25	1016	(CT858)	CT858 Hypothetical Protein
	1019	(CT860)	CT860 Hypothetical Protein
	1020	(CT861)	CT861 Hypothetical Protein
	1022	(CT863)	CT863 Hypothetical Protein
	1032	(CT373)	CT373 Hypothetical Protein
30	1033	(CT372)	CT372 Hypothetical Protein
	1034	(CT371)	CT371 Hypothetical Protein
	1057	(CT356)	CT356 Hypothetical Protein
	1058	(CT355)	CT355 Hypothetical Protein
	1061	(CT330)	CT330 Hypothetical Protein
35	1073	(CT371)	CT371 Hypothetical Protein

Coding Genes Not in *C. trachomatis*

	0486		Hypothetical Proline Permease
	0279		Possible ABC Transporter Permease Protein
40	0505		3-Methyladenine DNA Glycosylase
	0193	argR	Similarity to Arginine Repressor
	1041	bioA	Adenosylmethionine-8-Amino-7-Oxononanoate Aminotransferase
	1044	bioB	Biotin Synthase
	1042	bioD	Dethiobiotin synthetase
45	0585		Similarity to Cps IncA_2
	0562		CHLPS 43 kDa Protein Homolog_1
	0927		CHLPS 43 kDa Protein Homolog_2
	0928		CHLPS 43 kDa Protein Homolog_3
	0929		CHLPS 43 kDa Protein Homolog_4
50	1045		Conserved Hypothetical Membrane Protein
	0251		Conserved Hypothetical Protein
	0278		Conserved Outer Membrane Lipoprotein Protein
	0907		CutA-like Periplasmic Divalent Cation Tolerance Protein
	0171	guaA	GMP Synthase
55	0172	guaB	Inosine 5'-Monophosphate Dehydrogenase
	0608		Uridine 5'-Monophosphate Synthase
	0735		Uridine Kinase

	0980	Similar to <i>Saccharomyces cerevisiae</i> 52.9KDa Protein		
	0232	Similarity to 5'-Methylthioadenosine Nucleosidase		
	1046	Tryptophan Hydroxylase		
5	0477	yqeV_Bs	Conserved Hypothetical Protein	
	0048	yqfF-Bs	Conserved Hypothetical IM Protein	
	0587	yvyD_Bs	Conserved Hypothetical Protein	
	0143	yxjG_Bs_1	Conserved Hypothetical Protein	
	0448	yxjG_Bs_2	Conserved Hypothetical Protein	
10	0006	0180	0440	0977
	0007	0181	0455	0978
	0008	0190	0456	1018
	0009	0203	0457	1023
	0010	0204	0458	1027
	0011	0205	0459	1029
15	0012	0209	0460	1040
	0028	0210	0461	1051
	0029	0211	0462	1052
	0034	0212	0463	1053
20	0041	0213	0464	1054
	0042	0214	0465	1055
	0043	0215	0472	1056
	0044	0216	0473	1064
	0045	0218	0481	1065
	0046	0220	0483	1066
25	0047	0221	0492	1070
	0049	0222	0493	1071
	0050	0223	0494	1072
	0051	0224	0498	
30	0063	0225	0499	
	0064	0226	0516	
	0066	0233	0517	
	0067	0240	0523	
	0069	0241	0524	
35	0070	0242	0553	
	0099	0243	0574	
	0124	0266	0600	
	0125	0267	0656	
	0126	0268	0664	
40	0130	0277	0677	
	0131	0283	0678	
	0132	0284	0685	
	0142	0285	0686	
	0146	0287	0724	
45	0147	0352	0731	
	0155	0353	0745	
	0156	0354	0753	
	0157	0355	0794	
	0158	0356	0795	
50	0159	0357	0796	
	0162	0358	0797	
	0163	0365	0798	
	0164	0366	0799	
	0165	0367	0829	
	0166	0368	0830	
55	0167	0371	0831	
	0168	0372	0881	
	0169	0375	0882	

0170	0376	0913
0173	0391	0914
0174	0398	0930
0175	0404	0944
0177	0431	0964
0178	0432	0975
0179	0439	0976

1

Pn_0017
 mp_4 PM_1 (frame shift with 0016)
 LFTMGKGTWVWVDDVPTKNTLTKWYKYNDEPOTLVENHWEHNVQVRS
 LQMLPTKSHKSHNTLWVQVLTADTLEDDKINRQVTCACVAVXGFTFTSENFN
 ATYQLVYDKHNTLVAPNTHYVAVAMLVQVPTFLAKLHNSDNLVFNARFAYT
 TNNMTKTYKYVLEVCQWRLATVHGVGALVAVAXGRVWVTHFTFLNEMVYHQ
 DQNKTKHGRVPOFLNLTNLTAVHPTKREKDTYDLELVAVYVAVLRNDHQTTLQ

U0010 13324 14325
robust homolog present in Genbank/EMBL as of 11/7/98

VGGDQWTTCTGSLSRQALLVRAGNHFAACNFVEFQGFVEVLGGSSRSYAI DLGGRCFG

CpN_0018 27513 29003
pmp_5-Polymorphic Outer Membrane Protein
SYNMKTSVSMMLALICSGASSIVLHAATPLNPDGFI GEGNTNTFSPKSTTDAAGTYS
LTGVELYIDPGKGGITGTCFVETAGDGLTFLNGNNTLKFSLVDAGANIAVAHVQGSKNLS
FTDFGLSVITESPAGVTGKGLSVLGAVALQDINTLVLTSSNASVEDGGVIGKNSCLIQ
GKMSAIFQONTSSSKKGAIGTTOGLTENNLGTLPFNENKAVTSGGALDGAASTPTAN
HAAKFNKTDGASIVAGHIAAATHTHTTAAADNNTMLKQYAGATGTLITGSD
GTHVNTSGQGFALGAAMHIAATHTHTTAAADNNTMLKQYAGATGTLITGSD
GGDIVFEGQVTTTAPNATTKRNVILHSTAKMTGLAASQGNATFYDPTTNDTGASDN
LRINEVSANQKLSGSI VFSGERLSTAEATAENLTSRINQPVTLVEGSLVLKQGVTLITQG
FSQEPSTLLLDLGTSL

CpN_0019 29007 30356
pmp_5-PMP_5 (frame-shift with 0018)
ASTEDIVITNLSINADTIYGNKPINIVASAANKNITLTGTLALVNADGAFYENHTLQDSQ
DYSFVRLSGAGGTIIITQDASQKPLEVAPSRPHYGYQGHNVVPIGTGTQSOANLWV
RTGVLNPNRQGSVLVPSNLWGSFVDQRAIQEIMVNSQILCQERGVMGAGIANFLHRDKI
NEHGYRHSGVYLVGVGTHAFSDATINAAFCQLFSRDKDYVVSXNHTSYSGVVFLEDTL
EPRSPOGFVTSDSSEACCNQVVTIDMQLSYSHRNNDMKTYYTTYPEAQSWANDVGFLEF
GATYYYPNMTSLFDYYSFPLRLIQYHQAQEDFKETGGVREFTSGDLNLAIVIGKFE
RFSDCRGSYELTLAYVPDVIKDKSTATLASGATWSTHGNLNRQGLQLRLGNHCLIN
PGIEVFSHGAIELRGSSRNVINLGGKYRF

CpN_0020 32717 30603
Predicted OMP (leader (14) peptide: outer membrane)
KLWSPNRLMKRCFLFLASFLVMGSSADALTHQEAVKKNSYLSHFKSVSGVITIEDGV
LNIHNNLRQANKVYVENTVGGSLKLVAHGNVNVYRAKTLVCDYLEYEDTDSCLLTNG
RFAMYPVFLGGSMITLTPETIVRKYGIITSEGGPKDCLSGDYLEYSSDLSLIGKTL
RVCRIPILFLPFPFIMPMEIPKPPINFRGGTQGLFGSLYLGMSYSPISRKHFSTFFLDSF
FKHGVMGFNLHCSQKQVPEVNFNMKSYVYHRLADMAEAHRYRLHGDGFCFTHKHNVS
GEYHLSDSWETVADIFPNFMFLKNTGTTRVDCWINDYFEGYLTSSVKNVSQANQELP
YLTLRQYPIISYNTGVLENIVECGYLNAFSDHIVGENFSSRLAARPLKHTVPLPITG
TLLSTLSSLIYSDVPEISSRHSQAKLDYRFLHKKYIQRHIIIEFVTFITETR
PLAKNEDHYFISIQDAFSLNLLKAGIDTSLVSKTNRFRPRIHAKLWTHILSNTESKPT
PKTACELSLPFGKKNVTSLSDAEWIKKHCWDMNRWIEWGNDNVAMTLESRLSKYSL
IKCDRENFLDVRPIDQLDLSPLSKHNLILGLKFLVRHPHCWNYRLSLRYGWHRRQDTPN
YLEYQIMLGTKIFEHQWLVGVYERREADSRFFFLKLDKPKKPPF

CpN_0021 34470 32707
Predicted OMP (leader (19) peptide)
CGRSPYPIEILARGVEHRSMLFHLTLFGLLCSLPISLVAKFPESVGHKILYISTQST
QOALATYLLEALDAYGDHDFVLRKIGEDYLRKSIHSSDPQTRKSTIIAGLAGSSSEALDV
LSQAMETADPQOQLLVLSAVSGHLGKTSDDLLFKALASPYVIRLEAAYRLANKNTKVI
DHLHSFHLKPEETQCLSAIFRLRLETESDAYIRDLAAKSAIRSATALOIGEYQOKR
FLPFLRNLTSASPOQDEAILYALGKLKDGOSYPIALIKQLPDVTLLAAQAALIALGC
EEDAPLVKIKKALEERPRALYALRHLPEIGIPIALPIFLTKTNSAKLNAVALLELGC
DTPKLLVYITERLVQPHYNELALSFSGKRTLQWNRVNIIVQDQOERERLSTTRGLE
EQLLTFLRLPKAYLPCYIKLLASQKLOLATTAFSLSTSHQEQALDQLFAQALPGEF
IRAYADLAIYNLTDPKPKSLHDYAKKLIQETLLFVDTEHQPHSMPYLRYQVTPES
RYKGLDILETLATSKSSDRIALLQLMTEGDAKNFPVLAGLLIKIVE

CpN_0022 35042 34395
mal
TLIQVINSNCNVSNTRSFYSMSLPVLGSSSPRRKFI LKFRVPFTVIPSNFDESKVSYS
GDPIAYTCLAAQKAYAVSELHSPDCIILTGDTIVSYDGRIFTPQDKADAIQMLKTLR
NOTHDVVTISAVLHKKLLTGSETSQISLMTIPDRIIESYIDTGTLLNCGAYDVCHGGL
ILKKVHCVYVQGLPIQTLKYLLEENLIDLWDYSI

CpN_0023 36657 35014
yjaK/air-ABC Transporter Protein ATPase
ENRKLKYSKHOFVLMASIVLDKIKSGTLRILFDVSVVFNPGNCGYLTGPNAGKS
TLKKIIMCMIETPRGSI LKPKVGLRQNDISPHDTTVLDCVIMGNTRLWEALQRRDNL
LOEFTDAIGMELGEIEEIGEENGYRADSEAEELLTGIGIPNEMFDKGMAMIPIDLQFRV
LLCQALFGHPEALLLDEPTNHLDDYSINWGLNFKDYEGTVIVVSHDRHFLNTITTHIAD
IDYETIIIVPGNYDDMVEMKTASREKADIKSKKKKSQKFEVAKFGAGSRASQVQR
RETKKLQQLQGLKKNIQRPYRPLSDKSSGKVLVLEAITKDYGDHGVHFFSLIYQ
GDKLQIIGNNGLGKTLMLKLAVGAPSSGSIKLGHQACISYFPQNHSDVLADQOQETFL
EWLNRKRTGINDQEI RSVLGLMKFGGDADFKQIALSGGETARLLMAGMLENNHLVILD
EANNHLDLESVSALSWAINDYKGTAFVSHDRGLIQDCATKLLIFDKDKITFFDGTMDVY
TAGHKQLL

CpN_0024 37605 36661
xerC-Integrase/recombinase
REVMIASIYSLDYLMKVSASPHLRLNYCLDNLGKIFLEERCNLA PSSLQOLATEKRK
VSELPLFLPTKEHVRMYIAKLENGKAKRTIKRCLSSIKSFAHYCVIQLIKLENPAETIH
GPRLPKELSPMTYACQVEVLMATPDISKYHGLDRCLMELFYSGLRISIVAVNMQDQ
LSTHLIRIRGKGGKRIIPVTSNAIQWIIYLHNDRKRLLEKQPOAIFLNRGRRISTRS
IDRSFQEYLRSSQLSGHTIPTHTIRHTIATHWLESQMDLKTIALLLGSSLETTVYTVQS
VKLKKQTHQEAHPHA

CpN_0025 38610 37684
elac/atsA-Sulphohydrolase/Glycosulfatase
ILMSSRELIILGCSQQPTTRNQGYLFRWNGEGLLFDPEGTQRQFIFANIAPTTVNR
IPVSHFHGDHCLGLGSMMLRLNLDKVSHPICHYYPASGKYFDRLYGTIYHETIQVVEH
PISEEGIVEDFGSFRIEAQLRQHQVDTLQWRITPEDTIFKLPELESRGIRGLIIQDLIR
DQELISIGSTVYLSVSVYRKGDIAIADITLPCQAAIDLAKNSCHMLCESTYLEQHRHL
AESHFHMTAKQAATLAKRAATQKLIITHTFARYLNLDDFYKEASAVFPNVQVAQYRSYP
FPKNPLLNK

CpN_0026 39637 38762
CT345 hypothetical protein
INFAMGHIIPKLGKGVTSYFHKPQKQAAQPKITRQNTIAYLILICVLTVT/VLQAML
CMFTFVGLIPGLCLHALLVLLSILNPLCNINNTSPKYKEIAKDAJESQPTKASRKG
IQALSHHIDIRKNEIRTOLEKQVNYVNTKPKKTEGSPITDEHIIHQKPKTESSSE
CEHILHYPKAKAPHTATTEKSKNTTDLKKKKKTKIKLARTDSIRGAPKPMVPSK
RRKIVLKKVPLPTEDELEQKSNESGDSIPPIVQKALIFWYKQPTDF

CpN_0027 42752 39778 39778
Ion Ion ATP dependent Protease

PSIRTIIVDSTTNSDSP:LDNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQSSINQKIEATITKSQKEFFLKEQLKTKIKKELGLEKEDRAIDIEKFSR
LRKRHVPOYAMEVITQDEIKLQTLTSSAEYTVCRNYLDWLTIPWQIQSKEYHDLKAE
IVLNKDHVGLDEIKQRIELISVGLSKSLKAGSICLVGPPGVKVTSGRSIAKVLHRKF
FRFSVGGMRDEAEIKGHRPTYIGAMPKGMVQALKQSQAMNPVIMIDEVDKIGASYHGDPA
KILVNFNFTVEKLLDESEEEEDQSTERLLPSELFILPLNKRPF
PCMAAPILIESGPPYEVVLKVLAKSIOKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPICGSSAQVLLSIEERIRIIEPIKDKYLKARVSYHADNKELEELKAYSISIVSVIKI
LUKLINLPKEELQIFLCHSDFTPEGLADFGVALTTATREELQEVLETTNMHDIRIDKALI
LUKKELDLSRLQ

CpN_0038 52119 53811
CT318 hypothetical protein
MDTQSTGNEEDWIACTGIVGSMALGKVFVLTGSPHVRLETLPEQEEVEHRYKALN
RGSKDIIVALEBOEVVQOGLQEVGSLQALHLEIMKDPPLTEEVVNTIRKDRKNAEYVFSV
MOKIIEESLIVAVRGMPTVDRVQDTHDSINRVLGHCCQHQHSLGESQNLIIIFSEELTPS
FVASANTAYIRGFVLTGVAATHTAIVSRASIPYLANIISEELWNIAPKPNKLVLDGY
RIRWLLDYVILEDQQAIAKASLOQSIKVLIPGVSDVSEIEVKKWETIRTRFKPHG
KVSGWTMIEFPASVAMMIEILPECDFLSICITNDLVQYTLGISRESALPKHLNVTLPRAVI
RMIIHVLAQAQKQVQVPSICGAAQGLSLTFLPIGLGVQELSVAMPVINRLRNHIALLEL
NSCLEITEALLQAKTCEVEELNRNKKITS

CpN_0039 54256 53963
CT339 hypothetical protein
ISMGSYAKKKEAKIMEQQLFLEASLEKRYEQAGNGLVSVVINGKCDLISVKVQPT
CLDPEDFEVIDELFRAAFKLAKEQMDQMSLMRSTMPF

CpN_0040 55673 54318
dnaX-DNA Pol III Gamma and Tau
AFYTHSLGTYMTLPQYQASSRKYRPIREILQGSSVAVLKNALVFNRAAHAYLFSGIR
CTGKTKTLARILAKALNCVHLSDEGPCNQCSFCKEIASGSSLDVLEIDGASHRGIEDIRQ
INETVLTFPVKAKFKIYIIDEVHMLTKEAFNALKTLLEPPQHVFFATTEIHKIPGTI
LSRCQKMHQRIPEKTIILEKLSLMAQDDHIEASQALAPARAAQGSRLDAESLYDVVIS
LFPKSLSPDTVAQALGFASQDSRLTLDNALQRYATALGIVTDFLNSGVAFVTFHDLTL
LFYRNLLLTNSTTSKFSQYKTEQLLEIIDQLGESAKHLQNTIFEQTFLETVIIHRIY
QRPVLSLEISISKRFQEGRLNKTEPTLITQVSAPOQPTYKEQSFLEKKNQPAAGKII
SVEVSSASIKSAAVDTLLQFAVVEFSGLIRQ

CpN_0041 55888 57342
No robust homolog present in Genebank/EMBL as of 11/7/98
CKYLVHHSYPPQHSVSGSISSRYKRLVLAITFLVLGVLIIISGALFLTLGIPGTAGVSF
GLGIGLSALGVVLSGLLCLLVKREKVCPEEIPAVQPEETPEGVPTPEKALDEA
QKEQTKQKILQDLQELDQLDRYIQEAFACGLPKLDKLYEDQGLQDVKEEFQVDFVQK
DMIAEFVELQRIQCEGRLLLEFVQNTYIRGDLFKREDLYKLWELGLYLPSPDRYEPH
LKKSAREVVDLFMRRTTCNIRKIAMTFDRHVSVAKTAFAKAFGALETGVYESMRESYREA
PCEYKAKLLDGEKSAHAHQRFQDQKRWEDVDKAFVWKEGDKIEIDDAIQNSCKEKL
RYEHRITRARRYKVAEHQFNATRVKQSLREHNEARVAFEKERSKNEARQVKKKEKR
LRDQELHDLQELPRAQERLRELQALYPEIAVSVVEARREVASDLEKAHESIDKHQSCVR
EDEL

CpN_0042 57346 58182
No robust homolog present in Genebank/EMBL as of 11/7/98
EBEERKQAEFFRENGTKIRSMEEVSEYVQVENQLESCSKRLTKMETFALGVLEAKEEIE
SEESDVVNRFEVLCDRIEDMLSRVEIERMLMAELPLPIKEALTAKFVQNSCKEKL
TKWEPYKFEVLTSEERLQSLNCTQRAYKESQVSGLESEVRACREQLKQVQFET
QVSLIKEELFVTSTFRTKFSYHSFLHVPCKMRLYEYVDDIDLERTRARWAMSEYR
DAQAFQEMKLEGLVEBAQALRETEWLYREERKSKKKH

CpN_0043 58432 60372
No robust homolog present in Genebank/EMBL as of 11/7/98
HREFMQVPLSPQLPPPPDHVSQASCLSKFVRLAITFLVLGVLIIISGALFLTLGIGS
VSLGVLGSLALGSLVLSISGLLLELREVRVSGLEGIPTGIPVGPSEAPSSSEIQQKQK
AQKILDLQELQDLQDLDIQLHVLSCGLKLDLCKDRGLLDAKEKLVQVDFVWKDMME
FVELQVMDQESRYLEGLIHEVQSIHKLFDVDDVNIIRSHLGESCGYLSSEDVRGELLKRF
AKVAVRVMKVRTDIRKIAMAFNKNAAGAANKAFDQFQSGLETCLYKSLTKSVYDTFCDY
KRAKILPDEHNSARAQRFREVKDHEDLKETVFWKEDGRIDIEVLTAAGWGPDRYPHE
LLEKRRKDVMSHQLWEATMRVKEAEVTVSVARVAFEDKDSQOQKQKFKTERLRLCKL
DLRQCECHRAQERLEKLTALYPEVSVSVETTERERKFNLEKAYGNLEERYQSVQDQEDY
WTEOKNREAFRAKQTKVRSMEVVAEHQILENLELDCYKLSKAETFALGVEREATEEIE
EYTLSDAENRLKVLCEIDEDTLPRVEIEMLLMAERPLPIKQAFKAFVQNRCKER
LAKVEPYKESPAYVNSEERLQSLDQASQCIQRPVKGFKFRNGSMYI

CpN_0044 60278 60778
No robust homolog present in Genebank/EMBL as of 11/7/98
TAKSDCRVIRLHSAKESQVSSLEACTYREYLRQVQVQFETQVSLIKEELLFLSS
ILKSLSDPLIANIPCKMFYQYQDDIDKARAQSRWLEKSEYRNARQFQIYKGLF
KEAKPLKKEEYRLAQEERSNKKRLIYNKMAERQVQEFESMEIPE

CpN_0045 60961 62790
CT345 hypothetical protein
CKYTYHPPQLPPDHVSQATSWQPKRLITITFLVLGVLIIISGALFLTLGVPGLAAGLSF
GLGIGLSALGVLVLSGLLFLIRRGVSKVRPEEIPVTPSHEAQKILCQLQELDQDLS
IQEVVSCGLKLDLYEDQGLLTVQEKLRVDFVRKDMVTEFLQOQVVAQEGQFLDYL
INQVQSIHKLFPVDPVIGAHLAELCGYLPSCDVRVERLKRARQVDRFMRVTCDTKVN
AMAFDENACQVAKNAFKAQALGEECVYKSLTESYPEAFYVEKAKILRNEDEVWLQDN
KSARAEQRFREVKDRWEDLKETVFWKENGCTDLEVLTAQGWDRGPEHLIPEKRRNKV
MSHKLWEATMRKGAEGTYSVARVAFEDKDSRKNQKQYQEKTKWLRCLKDLHDOECHRA
RERLAEALYPEVSVSVVETTEREFKFLTAQYGNLEERYQSVVQDQEDYKWEENKEAE
FREKQTKVRSPEEVVEYLQILENLESDCSKQTLIAE//VLCVELEATAFEYTLISDAAN
RLKVLCEIDEDILPRVEIEIMLRIAELPFLPIKQAFKAFGLVNSCKDLAKVEPYQCE
SVDYKSGFRV

CpN_0046 62775 63263
No robust homolog present in Genebank/EMBL as of 11/7/98
BRFQSLNQDLQNVYQEQKATGLSEVSAYRDHLREQTFETFGQLDVIKEELLFVSSTL
KSKLDYDPLIADICMKFYEEYDGLDKARQASRWLEKSEYRNARQFQIYKGLF
DQALKKAEYRLREKRMNKKELIICNKEIAAQVQVQEGPQDS

CpN_0047 63437 63652
No robust homolog present in Genebank/EMBL as of 11/7/98
RHFHLKVTIVGFPRIVLEGLITMYHQPCKIPMTLTPGVLNKKLRKDYELWVYQSCPEG
KVLQYTSHEWL

CpN_0048 63687 65801
*ytdF is conserved hypothetical IM protein
MKELPHQVNRALAKILKSHQWVRVFLVTEVQCVIVAIFTFAWLVLYVPEYKAGEISRI
LTFAMDFSLWELAKIYKTAHLEAPKVVYILITLHGLSKKRNADENTDYWPKAAD
PILITNFDVSDQKLDLITPPLGSKKKTLEIINIKNKVIAQFCHLKLFLIQEN

CPQPCFDAIMDILKIANFEVAVDKEMSGCVKGLLGRKRCIEKITKTPILEKYQRIIDRD
AKILQRLAQLLGVHTLFSGRSLGWAIFVLLILLWGYGALKALCPPEMLKSPQRFMLYIA
LITLGLSLQCPGTEIFCAYWVYLYSYPPLPTAVLLGYFLGLPIAGFSCTFLALLYTLGS
DLWNSWFLSINLLCSWRILVSLHRVSRLLSVFWACMKLGGVAMGSLMFRIFTNTIS
ALYADGIESFVYSLITAISWALIPVFEAGPGATNFSLLTYLSPENALLKRLFKAPGT
YQHSVLVGSALAEAAQAIGADSLYCLVAHYHDIGKLINPGFFSEKQILQSGHSLSLPL
ECAKMIMRHIPGVNLRQAQGLPESFICQVEEHGTSVIRSAVYGHMVNPSGTFDEEL
FRYSGNPKSSKETTIIIMADSFEAASRLKNASLPDLQRLIDQIIQGLQDQGFSCSPIT
IDMLLTKMKGKILNHPMYKDEYQVAMRDEKRSQGT

CpN_0049 66362 66817
No robust homolog present in Genebank/EMBL as of 11/7/98
LKEKRRNIVYLLVIYQEIFWLTMLHQPYDKILTGNTIYIPGHTHKDSNKLFOKKSRAIW
VDEKPFSLDCFSVNLIFVSLVPIAGLVRAVYIKKSLDRITVQIGYSPSLSCDEKCEVA
FVNGYGLICISILGGLILVPLILVLVLSLLLGILMLFSLSTYESIKNYISKHICWKS
AT

CpN_0050 66849 66499
No robust homolog present in Genebank/EMBL as of 11/7/98
VSWFPIGLIFLAMYAKHQTWNDENVKANGLVPSSTNCKNALCRMSSRLTSSTIKTAGIL
GGCGILLPILLLLAILLSVLPQLIMLPFLRCCFALRQSVSSDVTNLLLNNTLA

CpN_0051 66797 67111
No robust homolog present in Genebank/EMBL as of 11/7/98
CFAYLIARNIPRMGNHETYIHPGLVPSHAQDVSRSTVYPSRSFIMRMLMGWNNFRVPS
KSSEQLMDGHRILPFFGKHHTTISILNVRNFWLSIFYNRGERF

CpN_0052 68008 67304
hemC-Porphobilinogen Deaminase
KMLSVCVSDCLSDFCQGRPLRIASRNSNLAKAQVHECISLLRSWPKLWFQSTTETT
GDREKKILPLHLVENSFFTDGVDALVHKGVCDLAIHSAKDLPETPSLPVVAITRCLHPAD
LIVYADHYVHEPLSPRLGSSSLRRSAVLKQLFGQGLDITRGTIEERLDQHRGHYDA
IVLAKASRLRLHLHAYSIELPPPYHALQSLAITAKDHAGWKQLFTPIHCHSS

CpN_0053 69350 67986
sms-Sms Protein
IRMATKTKTQWTCQCGATAPKWLQCGPCGCHNWSLVEEYVQARSSTSSRSSTSAIALS
SIELENERIFIDHAGWDRILGCGVVRGSLTLGGDPIGKSTLLQTAERLASQYKVL
LVSGEESVQCTSLRAKRLNISSPLIYLFPETNLDNKKQIATLEPDLIIDSQIIFNPT
YNSAPGSGVQREVTYELMQAKSAQITTYEIGHVTKSGEAGPRVLEHLVDTVLYFGN
SHANYRMRISVKNRFGPTNELLILSMHADGLKEVSNPGLFQKQGTPTTSGMIFPIEG
SGALLIELQALVSSSPFANPVRKTAGFPNRFSLLLAVLEKRAQVLFMTDVLFTITGGL
KIIEPADLQALLAVASSLYNRLPNNSIVIGEVGLGEIRHVAHLERRIKEGKLMGFEG
AILPEQGISLPEIRENFRLOQVKTIDAIRLL

CpN_0054 70089 69313
rnc-Ribonuclease III
TLSPFPPIKIPNSKFKDGALLSMHPIDITAIEAKLNTFTQPKLIEALTHPSYKNESA
VQIEDSERLEFLGDAVLGLIVTEHLFLLFSPMDEGTLSTARASLVNAKACRYTMTLGIG
DYLLIGKEIKQSERGLSAYANLFEILGAVLYDGLSPARKLTVPLPPEILPLMS
GNPKVLLQQTQKQFRLVLYQSTAVTDAQGNVSYQIQLVNVQEWEGNASSKKEAEKI
AAQALDITYGNQNTMDV

CpN_0055 70096 70590
CT296 hypothetical protein
CFWICYLIRIRMSALHLQHLRHFHNGHSILFENLLTKDCFLLETKLQNFIAKASKTID
TVWRNENIFRSMPEITYTVVRKRLDFFAELVHRPKSLVRDLWVFPGEIELEGEDCML
FLLSGDRAGSGIIFTGPYPSDLYELEKGTGLLLAFSSVGIPIVI

CpN_0056 70917 72746
mrsA-Phosphomannomutase
EFLKLSLHRIISLMKEVEQIRSLSYDAVTAENICRWLSDNCTQDAKTLGLWLDTPAQLE
DLFGATLTFTGTGSLMIGITNRINLPTIRRTTQGLVQVLRALPHQDPMRVVVCDDT
RHNSIEFAQETAQVLGNGCEVLFQYPEPLALVSFTVYVRAIGGMVITASHNPNYNG
YKVMASGQVLPDLQEIQAACSAVNEILSVPSIDHFNHILIGKEYEALRDLTKQLQL
YPEANRISGRSLISISYPLHGTGSLVSLHDKWGLFSLVLEKQAGIDGDFPTVQLPNP
EDPEALTGLIEQLANDDLFIATDPAADRGVVGLDQGPYRPNQGMASLADHILGA
WSKTRHIGELKLVKSLVTTTEMLSAIAKHVYHVDLINVGTGFKYIGEKTESWRNSTKVF
GAESYGCYLYGTHVEDKDAIIASALIAEAALQQLQKTLCDALLSLYETVYGFANKTES
VVFSAKDEQIRKLLSHLEETISSANFFSGKYQVEKFENYKQIGFNLLSKDSYALTLPK
TDFPKFNSNL

CpN_0057 72913 73554
sodM-Superoxide Dismutase (Mn)
ILKRYVMSFVYSLPELPHYDVALEPVISSEIMILHHQKHQIYINNLAALKRLDAAE
TQONLELIALPALRFNGGHHINSLFWETLAPIDQGOQPKHELLSLIERFWGTMDN
FLKILIEVAAGVQSGWAWLGCPCAKQELVLQATANQDPLETLGKPLLGVDVWEHAY
LQYKNVPMYDLKAFQIINWGHINRFSSEISSK

CpN_0058 73627 74562
accD-AcCoA Carboxylase/Transferase Beta
IRWLRLFSYDKPKIKVQIKADGFSGLKCNHCHEMIHANELGQNYNCCPKCSYHYRIT
ATERYKLLADKDSWRPLTYDLKSQDPLEIDTDTYANRLEKARKNTTESQVIGICTIG
LHPVALAVMDFNFMAGSMGAVVGEKLTPLIEEAIETRLPVIIVSASGGARMQESVFLMQ
MYKTSAAALAKLHEAGLPYISVLTNPTSGVTASFAALGDIIEAEKALICFAGPRVAVQ
IGEDLPEGAQSEFLLEHGMIDKIVERKELKTLQTLDDYFLAQEYTGKSKAPRDLSCR
LPEIFLLTDDSE

CpN_0059 74562 75050
dur-DUTP Nucleotidohydrolase
IKHTATACNDNIICNALMTVFCLEDDQGLPEVTTTGAAGADLRANIEEPJALLIYQRA
LHTPIKAEIPEYGLQVRDRSGALYHGTILVINSPTCDSQYRIRVILINRDIJTFI
IEPKMFIQVVLSPVQATFVVKQESLAETARGCQGFQHTGAS

CpN_0060 75004 75528
pdsN-PTS IIA Protein
PKLPEEYEVVLIDQAKMPSYQNOQDPSLPSLLSPRLVMPGLKHSRDEITQDITDIA
AGLLEDQAFQADALVRNIMDTGIGWVAIPHQKLECCNFFIAGIHTQILWADIG
ALVRLVFLIGGPENAAQYIKLLTFLTLQIRECPQQLQVNTIEEMNVFVGM

CPn_0062 76251 77690
 EVVVKPKQKSGKQAKKEPRARKVLVPSSTLTSGARAKMKNSSRKESGQCNESISANST
 PSVKLLRNKRKAQKAAKQGFANSLTLKSLLPKLQKDNKSHIREKATSRFVNESQL
 SSARKRYCTPSSAAPSLFLETEIVAPVRLTEPKDSEIHPVVOVQTNPKQKNTTKTQ
 LQAQASIQSGTEKVLRLAQLAAGSLPVLVRNPNPVSVRQKEELLKELVAERKQCKRRK
 VQALAEARSLTEKQVARGGVSSTLTLYRDEKAAEIKSRNNCKVSPEARQKVSCKRDARA
 NGKQDKTTPSEDASEEQQTGAGLVKTPKQSVAGNAQFNRYRSKNTNIDSYLTANQVSC
 NSEDTWPCSSVSKVSKRNTSIVCTVMVTJAMIGALIIANATESSTSDPTPTTP

CPn_0064 78340 78576
No robust homology present in Genebank/EMBL as of 11/7/98
LVMTKIQCSCAQQYRSRPAERAQTTPQPFLARDRADFWERHPRFSACCRVLLLVAVVVLAL
LFLFVMLLPLAAGSYLLAF

CpN_0065 78882 80651
CT288 hypophthalmal protein
YDYKYKFFPKFKNYMTDFPTHFKGPKPLNKIVKNPFFERNPKVARVLGITAVVLGIALL
SGVLILITPLGPAISIMILGDLGSLAGGALFVGGTITATILQARNYSKAVKNQKLSLEPLM
EPELTKALDYLDLKEVDEMDLHGVSHKHLKLLDLSKTQREVLNQIKIDDEGSLGCEAA
MISENYDACLKMLAYREALLKEQTQYQTRFNQNLTHRNKVLILSLIRITTENISKAGGVF
SLKFTSLSSMRSIHTMTTIVILASVVSVMVAALIPGGLIALLVAISAGSVITWG
LSYLVRQILSNTRKRNQDFYKDFVKVNDIELLNQTVTLQRFLFEMLKGVLKEEVEVSLG
QDWYQYITINAPIEKRLIEIRVITYIEIDAQTKMKMTDLFELENNVSRGLSVASPSDEP
SETPTIPFGQKEAPKLRQTSQNTSTYTGPDNENIDPEFLPMPWMPKKEEDHSLSEPVTKL
EPGSRREELLLEVGNPTIRELNRNKLALQQQLSSVSRVKRHRPGEHYGNVYSTDLELDRIQ
MCEGAFTHLLREACEETIPOSGLDVLIDNRILGILVEGDSLSRTEEEPOE

CPN_0066 80916 82655
No robust homolog present in Genbank/EMBL as of 11/7/98
GPMANPTQSRPPSPSEISIEELQLLAGSNTETISMTPPSCAAATAEVSFLTIEGGRR
NSEDEEPIGLGSCVEVDFVITNNGDPEVRDHEVRVMYINGSGRTQHEGILDAMNICTLGR
FVDFPIHNSGQVCGSCVLLNIPRPNRVISQAIQARVNEFFIAENANDRYIVLFSGN
CGLYLQVALDNLVSHYSHILTCVGISSSYIQQGYRVNHYRVTDGNTLLDRGATATVNTT
LEPDSADAGELFLPSVRCPSYQWALRGCEQLCLNQQVGRFQDSSSEIALVNNQDH
SEVTRLIEWIDRGDSQAVLELNNPQSHCRDIALTALYATTRISSLGLECLMSVTYAPEV
FPAFAIVTQSYIMTLRYVILLTNRCGRRHFLRLAALQGLSGFLVTLGLDINVTNR
VNRPRRLVSVICTAFSGATFSGFIYVDLTRMFFTSRLRLQGLVQRRLTGRGLPLRRVFN
VGLDLRFQSQALITFHGFMPLILIGFNQLVYIQVPRVIRPNTTAYVDLNTQSGEAWDS
GVLDSLAGQTINFLCMLLVITNTFFVSVVRNLNRRHRR

PFE_0067 82920 84053
 NC robust homolog present in Genebank/EMBL as of 11/7/98
 GSSGYSGRPPMAVEGRVNSSQALNQDCEVLANKQSGKGLRCRLISIVVAITFIAGVV
 IALTILASLTISVPYLVAGVLLIFLITLQCIIFALCSKEIKHVPPTTISKEEIIATPEER
 GEMKEKEKEDGPEHFHGTATDIPMSRALQDQFNHSEHHKISPALETYRSHQDVLLFKDW
 PVLPDVTSEEEVLIRVSVGYSLLEACVPKVSMLIDELNKLKSPREHCDFIDKTKL
 CPTSLFTQKDLTFFLAYTRVNDGHLAPFRAGAKVLIHYVRLRQHQNCDFTPGHS
 YCARLAFNQTORLYLQHLFVNEKLRSIYANMKDKDPLCHFPWAFIPIYDLTKTSDHGDGFLE
 QEDREYPSRAQDOHFG

CP1_0068 84909 84331
T360 hypothetical protein
PF1HKFFYILSIFCSGSAPLKGCINEDVSSQSRIEEDPEVLITQLNELIETPIEGKE
RELQAIISDGQKSSEETESCGTSDSEGLSEKTDKESSENYLVDFDSMQRLLEGISK
QCSQVAQIIDCFNREFDIRNRELKLNRELEKREKDLFKKSIIDWNKEKVSRELAQR
DQIDKITMLLKK

Pn_0069 85191 87086
 No robust homolog present in GeneBank/EMBL as of 11/7/98
 NFLVLYLFLNGLIMTPPPSRSSSSPPGDIWIELQDLQDGLNTNNSSRATPPPPVEVGGLP
 YFSASNFVVIWEDGASLPSPQQLSLPEYSRQPPGTVDETASITRTSEEMPTLVST
 CPANSEDRNIEDEHVNCLYIASTSDQLAEVQGGMHITLGECPVRYLYETGHLAYAFR
 NTHSHRLEVSHTVRAMTYPWDRFFSRHNWNGRRFLVFYQNGQVGAQALDDSSMTQDI
 VLGLSPSTVYIRGNHYHQHYRVGFQWSPCLDLSAACAEVSLPYGESSDGIYPSFLFSH
 FDNRIYRGERCLTSGEGMGLPETQQYSLTSLLEGHAEVLVNLQNPQAEIASR
 MHNEGRGRLSNYPGRSSNPFMTSQYVLRNLNTLGAIIYLPSSYFSQSDNIVCLTIS
 AAETVLYIFLTVTSDTCGRRLYLRVPLCLTGLRNLALPTLLELLLSYRSPVEGVP
 RAVETGLVGYMCTTRRVFAWNLILRWPRFCRLRHGILQVFRHSITGHTLGARITDLTLSR
 ATVPFSYSGCSSLTALAHANTNLTALDPYRILTESGLDRLRPAFNDDEMQQADNPWDAYS
 IVLNTCITYMLILFANLIFMYVSVRRYHRR

Pr_0070 87399 87208
 α robust homolog present in Genbank/EMBL as of 11/7/98
 KVGLFHLKKNQFFSNQSRTYEQRFPKVSPIFESILPLQSVGFSSQGTLLISFRDTLKR
 LYI

En_0071 88066 87599
F125 hypothetical protein
KSLRLLEFICPLQHARCLKKQHKIIEELFPEPPQGDHLYLKLMMENSSSRDAFDKKRML
ENHIVQQQGLDLYLYEYVQDGLFFETTYTKALMSGIAGLFTEVYSGSTPTSLTCKKIF
GRLTFLPGLRLGIESLYMRMKGIAVQGLKPGCT

Trn₁-0072 89151 89057
 Y124 hypothetical protein
 Y124:FTKTVYKEKVLITLYCLDPFFHYHMYMTPLCGGICSPSDQYVQELFCDLSSSR
 NPEPCNAKXGGGIVPPILALVALTDLKLVPYNGISFSTWTRLRNAVEKICGLFLQRNWK
 IITLITLAWALIVCHNITVALTLLTWLGVCLGIGVPGIPATATCLDKENKRRHVNLSWNL
 NHGTLIDPPLTRQILLATNITASTALIVAVPQA/CLVIGFSIGNQLSINTVYGAALGD

CpN_0073 89353 89574
infA-Initiation Factor IF-1
SMAKKEDTLVLEGGVEELLPGMHFRVILENGMPVTAHLGCKMKRMGNIRLLVGDRVTVEMS
AYDLTKARVVYRHR

```
tuA-tublongation Factor Tu
EDFEMSKTEQFQRNKPNIIGTIGHVDHGKTTLTAAITRALSGDGLASFRDYSIDNTPPEE
KARGITINASHVEYDETPNRHYAHVDCPGHADYVNDNITGAQMMDGAILVVSATDGPAMQ
KEHILLARQGVGPVIVVELFNKMDISQEDALIKDVLMELESLSEEEKYKCPSTIRGSAOL
KALEGDANYIEKVELRMQAVDDNITPTPEREIDKPFLLMPEIDVFSISGRGTVTWGTGRIER
IKVSDQVQLVGLGETKETITVTGVEEMFRKEPELRAGENVGLLLRGIGINDVERGMMVQCP
NSVKPDLGKFSAYVVLQKEGGRRKPFSSGYGPPFRFTDVTGVTDLPEGTMMVMMPGDN
VELMDVLIGTLVAGMFRFAIRECGRTIGAPISPKINA
```

CPr_0075 91087 91350
secE-preprotein translocase
SRSWFMKQQHNRKALSRKIGTVKKQAKFAGSFLDEIKKIEWSKHDLKKYIKVVLISIFG
FGFAIYFVDLVLRSITCLDGIITTFLEG

Cpn_0076 91334 91903
nusG-Transcriptional Antitermination
QFQCIVNCKMYKYVVQVETAQEKVKKALEDFKESSGMTDFIQEIIPIENVMEVKKGEH
KVVEKYIWPQYLLVKKHLTDESWLYVKSTAGIVEFLGGGVPAVSEDEVRSILTDIEKK
SGVQKHQFEVGSRVKINDGVFVNFIMGVSEVFDKGRLSVMVSI FGRETRVDDLEFQV
EVDAPGQSE

CPN_0077 91956 92435
rll-1L1 Ribosomal Protein
FFVSYPLFVEVSGCKVRFMSVKKVIKI IKLQIPGGKANPAPPIGPALGAAGVNIMGFCK
EFNAATQDKPGDLLPVVITVYADKTFITFKQPVSSLIKKTLNLNLSGKIPNRRNKVGL
TQAGVEIAIEOKMKMDIVLLESAKRMVDEARSMGIDVE

CPN_0078 92453 93160
 L1-L1 Ribosomal Protein
 SCRLMTKHGKRIRGILKNYDFSKSYSLREADILDKCQPPVRFDTVDVSIKLGIDPKKSD
 QQIRGAVFLPNGTGKTLRLVFAASNKVKEAVEAGDFMGSDDLVEKIKSGWLEFDVAVA
 TPDMMREVGKLGKVLGPNRLMPTPKTGTVITDVAKAISELKRGKIEFADKAGVCNVGVG
 LKLFSSQSQKENTIEALSSALIKAKPPKAAQQYLVSFSTISSTMGPGISITDRELMS

CPn_0079 93170 93688
r110-L10 Ribosomal Protein
RGKMKQKETLLLLQEVEDKISAAQGFILLRRLFTAAYSREFRNSLSGVSAEAEVLKKRIF
PKATEAAGLELVDCSDTDGHLGVFVSCGDPVSAAQVLDLFNKQHKDLSLVFLAGRMGNASLS
EATIEAAVLKPSKELRGQVGVGLFAAPMSVAGVIGMNSVLSGVI SCVDQKAGN

CPn_0080 93720 94121
F17-L7/L12 Ribosomal Protein
VRVTKVTTESLETLEVKELSNLTIVLELSQLKKLLEEKWDVTASAPVVAAGGGGEAPVAA
CPTEFAVTLEDVPADKKIGVLKVVREVTGLALKEAKEMTEGLPKTVKTEKTSKSDAEDTVK
LQDAGAKASKFGL

29n_0081 94219 98016
 rpoB-RNA Polymerase Beta
 PIRILSHQNSRRITMLKPKPVRYSVKKEDIPDLPNLNIEIQIKSYKQFLQIGKLAEEERIN
 LKEEVREIFRIPKSYNEATVLEYLSYNLGVPGSPPEECIRNGITKSYTNLKVRFLLDTG
 LKEEVEYMGITPLMDTKGTFTIINGAERVYVQVHRSGPGINFGIEKHSVGLNLSFSRIIPY
 GWSGLEAIFDINDLYIHIDRRKKRIKLAITFTIRALGYSSDAIDIIEFFTIGESSLRSE
 DFALLVGRILADNIIDEASSLVYKGAGELSTAMLRMLKRLDASIVASIVADDENHPRI
 LMLKADPTDSYAELKDFYRRLRGPETAPLANARSTIMRFFDPPKRYNLRGKRYGLYNLRK
 GYSIDDEALSQVTLRKEDVIGALKYVILRLKMGDEKACVDIDHLNARVRVSVELIQNQ
 RSRGLARMEKIVERNMLDFSSDTITPGKVYVSAKLASLDFDFFGRSLQSFMDQNTNPV
 LETHKRRLSALGGGLNARGFEVDRVHSHYGRICPIETEPGPNGLITLSLSFSAKI
 FEAGTFTPYRIVRDIIVDEIVMTADVEEECVIAQASASLDEYNMTEPEVCWVRVAGE
 FEADSTSTVHMDVSPKQVLSVTGLPIEHLDDNARALMGSNNQRQAVPLLNTKEAPVVG
 GLECRAAKDSGAIVASEDGVDFVDGYKVVVAAKHNPTIKRTHLKKFLRSSNSGTCIN
 QPLCAVGDVITKGDVADGPTADGTRIDELAGKGNVLVAFMPWGYNFDAIDIISLEKIRD
 YTSYITIEEFLTARDTKLKEETIRLIPNVSDELANLGDGDIIRGAIEVAKDPILVGK
 PKESETELAPEERLLRATFGEKAADVDKASLTPVPGTEGVMDVKVFSRKRLRSKSDDE
 VAEAVHLKDLQKGYKNQVATLTKTYREKLGALLNEKAPAAIIHRTAEIVVHEGLFD
 IETRIETIEQEDLVLLMGNPKMEYVLKGLLSDYETALQRLNIYNTKEVHIRREGDADLV
 IRTQVYVYASRRKLQVGDCKMAGRHGNGSVIPEADMEYVLSNGETVQMILNLPDVG
 RRMNLQGLVETHLGYAATKAYIVYKTPHVGEGFQRIWDMMIIEQLPDEMGQFVLYDGKGT
 FDRFNKVVYIYMLKLSHLIADKIHRSIGPVSLSVTPQPLGKAGKCGQSGFSEMEVAL
 YGVAMHMLQELITLVKSSDDVSGRTRYIESIVKGENLLRSCTPESFNVLIKEMQGLGLDVR
 MVVDA

LOC_0092 97992 102221
 poc-RNA Polymerase Beta'
 SSYGRRTKRLNDVLEKIFMFGNSRDQTVLSKEGLFDKLEIGIASDITIRDKWSCGEIKKP
 TIVNRRLLTKPEKGGLFCEKIFGPTKDKKQGLKTKIKHKGIVCDGCGEVLTSKVRRE
 LAHELAVLPVLIWFFKTFPSRIGNVLGKTKACDLREYIEEYVYIDPGTKTKQLQLLN
 QYQREVVEVKHGGDAFKVSMKEAIVYDLLKQLQLDLKELRTRTKSQQAKRKLAKR
 KILITBGVSSNNHFGVWVYKNTVFYFDPDLRLVYLDGGRFATGSLNLDLRRVITNENRKL
 LILRLKTPVYVNRKMKLQVADALFDNCRHQRHVGPMQAGNRLKLSGLMCKNCKRFP
 DGRGAPDEVDSRGVITVQPELKNPQKLPKDMALELFPFTIKRLKLSGYSVYISAAKMM
 LDRGPAVEVDVDELKIKVHPLVLPNAPTHLPVITQAEFPVLEKARIRHPLFSAEENAD
 GQGMAMHVPLSVEVQLALNLMMAFNPFLFSPKSKPVAISKDMTILGLYLMADPTFY
 FPHGQTKIPKDETEVRLALNVYFIDHVPFPRDCTYRGTHIKIVRVIDQITETT
 RVNPLFNHVPKELGQFNYSMVYFCTLLVLPVTKKVCLEATVIRFDIDLDGFLQATKA
 LGMVLDKVRIFDKIKLTKDAVYVLPVTKVYDDITETEGEIRGHIITFVEVSGDEL
 YVYCIKQTRSKHNPLFLMIDGFAVANKSLPGLVALRGLMAKNFACIATETITNRE
 TLEVLYISGHCAKGLADLAKTKADYVITPLPGLADALQVITETKIKTEHLNLEISA
 GKEELLPLKDRIVGRTVAEDVYFGDKVITLAQSGDVNSVQAEALDDAGITETIKRE
 TITVPEWCAKGLGNLANPLGMHAYVITAAQSGEGFTGTPMTIRFPIVLAATSG
 TITLNSDGILVMDLRRVATGEGHINIVLPIVGLIIVHDEGRTNIEVSKLLZTKGIE
 ERLVETVCLVLAEDVDFVQVGLVGLVLIHPIIDCKPKIKYEDVLVITETKVV

HIWMEVLSLQSP

CPn_0094 113104 115971
valS-Valyl tRNA Synthetase

[illegible]

CPn_0095 115956 118790

pkcN-S/T Protein Kinase

ACIVCLDQEDQRSLEYDIVIRIIGKGMGVEVYLAYDPVCSRKVALKKIREDLAENPLLLK
RFLREARIAADLLHGVVPVVTYISEKDPVVTYMPYIEGYTLTKLLKSVMQKESLSKELA
EIKSVGAFSLIFPHKICCTIEYVHSRGLIKDPNDLKPDNILLGFSAEVILLDGWAAGVCEE
DLLDDIVDSKEEVLSSRMTPIGRIVGTDPYMAERLILGHPSKSTDYALGVGVLYQMLTSL
FPYRRKKGKIVLQDGRIPSPQEVAFYREIPPLFSAVVMRLMVAIDPQERYSVSTELKEDI
ESHKLKSGPKWTLTTALLPPKKSSWKLNEPILLSKYFPMLEVS PASWYSLAISNIESFSEM
RLLEYTSLSKGLNEGGLIPPTSENALGDGFGYGGFWLHKERTLSVSLVNSLSIEIQRS
QDLESKRTEFLSLADQHNHSLSLFVDGTMVLHMYLPSRSGVAIVTRMDEIDETIGI
FESSGSLRFLSCLAVDPADFAELKYLDRAVLVYRLIAESFGPRGKEYEARFRAGITVLEKAS
TNDNEQEFALAIIEFSLKHDGVAAPLEYLKGALVQRYQORVQENIESKLLALLKRSYQHP
IFRLKHQVYVLRHESFYSKGRDLRALVMFVILLLEIAPQAIPTQGEKIIVLWLKDKSRATLFC
LLDPTVLELRSSKMELFLSYWSGFIPLHNSLFHRAWQSDVRLAIEIFYVACDLHKWQFL
SSGCDIFPKESLDDQKATEEIFVSEDFDGLAFQISIFNKEDAKEAIFVSNDDPSLLVY
YIFDLFANRALLQESGAEATFQDLIRSKVPENFYHDYLRNIEITRAHLWRKNQALSTP
ENYTEQKLQDQEHQELFVLYGCVLILQGAEEAAQMHVYDCREDRIIPASLLARNYLRGLR
KDALSYQERRLLLRKFLFYFCHLGNHDERDLCTQDITTEEFL

CPn_0096 124347 118837

ETFLSILREFFMKS L P V Y V S G I K V R N L K N V S I H F N S E E I V L L T G V S G S G K S S I A F D T L Y A
A G R K R Y I S T L P F T F A T T I T L P N K P V E E I L G S P T I A K Q N H F S H Y S H A T V G S T F E L F S H
S T I A L T E G A R P K P K T E V I D Y S K E A C I S T M L S E C S T I A L T E G A R P K P K T E V I D Y S K E A C

LGFTTRVRCNGTTHPIYSFLTSGIPEDCSVDIVID
 CSVL SDEELMTFESTKQOQIDDTYTRI TOOL ESRH

NSVDEELMTPTSTROQIDDDVYTYLPTQLQFSPHALESRCSLCOGSGIF I ISDNPLIDE
 NLSIKENCSFAGNCSGNYHYTYLTYQALADANFNLETPWKDLSPETQNTFLRKNKNLVL
 VRLFDOTLQGNKNTLYKVRWGLNDIDGKRVYTPKPSRYLSKMGSHASCSLCOGTGLDYA
 VRLVATWEGKTYTFEPMQSLNHNMHVFSKVSFSPSLSEITQLLQGLKQRLSLIDLGLGYLTPN
 VRLALATLSGGEQETRAIAKHLGDLFGFTYILDEPSEIGLHPQDTEKSLIGVYIKKLRDQNTY
 VLEHEERMSIADR I IDTGGAGI FGGEVLENGKPFDEPNNTNGKLYARQLBTITP I
 SEENRAPTSWLLLTATEIHNKLNLSPLRLARIGI VYVSGGSKSLINNTLVPAIESFTK
 VSRNPKLTHFEMWCIETRLHNITDLPGRSQRSLPTVYKAFDDIRELFASQPSRLRQGLTK
 VYFSAFNQOGACQCOGLGTMTISDDTDPYSCGQCGKRYHREILYEGKNLFIADPM
 VYAEAEAKY I SHPKIEHKIHALCSLDLDYLPGRSLTSLGGEIQLKHALLELLFASPMQ
 VYLVLDPTKTLGHTIDQALVELSLTYLGHYTVLEIHNHVVKVCYDVELGPEGGDL
 VYGLJLACSTPGLHDIQLNTLPAKALPYIECSLDIPVVKSEPPSSVKCSDILIKDAYQNNL
 VYHIDLPRNSLIATAGPAGKSHSLVFDLILYASGNTAYAE LFPVYIROGQLLTPSPV
 YVEGLSPVYI SVRKCSNNSRYSHYI TASALGNSGLEKLFALIPGESPFLTEELKSTTPV
 Y I IDSLLSKYKDDYVYITSPYIPGSGOLEIFLQKEQKGF I KYISEGNYLDLDERLPLNI
 YPAIVIQHTKYVSPKNSSLLSAISVAFSLSEIWTY I SQKQRKLSYSLGMDKKGRLYP
 YTHQLSSLDGPGRCITCGGREGELK I LSEEHKEIAHYTLEFFSLFFPKYSPMPVKQ
 YKLDKENASQPKLLRTKEFLNFCRSGSEPPFNALLMEQDLETSDSP LITKPLLATSCPA
 YKGSGLNDYANRYVINTSLSDITLQDATABETFLSFLNTIGTDDTRSY I QDLNRLTP I SKV
 YLSTYTLQGRDITLSDGENYRLHAKK I ISLNTINLYVFEELPSGLHPDQLPTIQLKLE
 VANNNTYVITDRSCSLPHADHAI FLPGSGPQGGFLMDSDETVCPSPVDLHANVPOTEV
 YPKPALSYISKANHTGRSDRLVNLVSIHITQNLKVPALHVALIAGVYSGSGKTSLLEGG
 YKQKQAEILLAKGTTSDLYVVDSPH I ASSQSDISTYFIASFRLAFRYSITQAKNLI
 YLTFSTNTKQGCSDCQGLQWIDRAFYALEKRCPCTSGFR I YQALAEVLYEGKHFGL
 YLMTPTIETVRLPFGFK I KQYKQLKALIDGLGYLV I QGKLSLSVSEKTLATYFYLF
 YPETPTFLIDELFSLSDLPIDKQHLEPKRLSLINSGHSVIYIDHVVKLLKSADYLIETGP
 YSGKGGKLLFSGGSPDILYASKSLDLKKYICNEEDLS

Pn_0097 124549 126006

MK - Pyruvate Kinase
 SMITRTTKIKITIGPATNSENPEMLAKLLDAGMVARLNFSHSGSHETHQCAIGFLKELREYQ
 PVLAIMLIDGKYIEPRLRLGNTQPIQSVYQQQKLLRVLSVSDIDGSAEGCVSLYPKGIFFPVE
 LADLDIDGYIHAAVVSEASDISELEFMSGLLKHSHKSLGIVDVALPPIQEKDIDLK
 YGMRGMDVGLASVRYGDEIDETMRKCLKQNPCKMTAKTENLRGLVFNFSKIAKLADG
 IARGNDLGLLSVSEVFNPKQKMAKVSRETNQPKPMLEMSIIRNVLPTRAEVDSDY
 IYLDGSSAVMLSGDSEACAHPAAVYIMRSLILETEKHLNLSDFLKLDDNSNALSQVSPY
 SASIAGLIGITAEADAKALLVYTEGSGSPMFLSKYLRPGCIATVETPSVYRLALBEG
 YPMLETQESDRAVWRHQACIYGEIQGILNRYDIRILVLSRGKCMETNTNLTVLIVNDILGT
 FPEET

En 0098 127494 126071

robust homolog present in GenBank/EMBL as of 11/7/98

/GCKFQHQIKRTILEAPYLYVLGIGLILRITPSPFLYLYGKGGTFAFYLYDYRKRL
 LALAFPEKTFDERYKIAQROLOHLTGLLEALLAIDOLVNIIDKLTITVITSRPKQFS
 EVIIEDELECTFKNTKQSLQKILFLPGLLGNWILPELITKYNVIAFAKAIQNRKLS
 LALPVEVFKLVVPPKNTLQGLIEADYLRKILVITGALLAMLVYVITPLGSPAPTTTS
 LALVITGPFVAVNVNIRQAQGTVPVILAKLANKRLMPLVVAITLMQMLLEKGLA
 VITLWMLHWRKWRPKLVNVIKKYKVTILVITVYVITHTIFALAEAKVITLTHLAL
 ADHLLEPQVPEYGLZLNQDGLLNLVTPAILITLNLNGLYKHPKFTQSCAVY

DECLARATION OF INTEREST The authors have nothing to disclose.

nt_0099 127827 127866 127866
 ...cytokil homolog present in Genbank/FBIH, p. 1177798
 CACATGKLSRFAAKHAPI RMNIN LCTTPTTPAATLITLVKVIVCAPSTVQRMISIK
 TLAPAPSPADTVATEALDELSTQVPLIACAKITDQQLPHIKPILTKF

LLKTFEGLIAEELNWKNVLTDEAPFINTVTKPNFRMLGKKVGGKMKVEVQKALSELPNNA
AIDKLIQEETWVLTDIDPEIALGDDVVICPHDFDGYIARGALFSVILDCQLREPLIVE
GIAEELVNKINTHRNQQLEFVGGRIALRIKTTAEVHRAFLDYENYIEETLI IAYDFDQ
SDFOEENWDINGHATQETGSSIDS

[illegible]

CPn_0111 144761 143934
CT021 hypothetical protein
QLQNYRPIPMNDSSYFERILQKYLMMKGQKGLFLFLFLSFLSFSTAFSGGLFASQTSSSLRT
IQENIFLAKTGQDVTVLSRSGORTFVLVNSTPKPTVWEIITHPCIAHKREPSLEQASWKT
VHQLHESPLQGVFVVLSSSEGQSFLLNTRTKSLSEPVGKSTTVPAFLQIFDLPLSPAPANV
IKTKGKENGKWPSPKVSFEGLAPTSISVNAWGLNPKDRGPLSECTILMYTPQDISVFFL
WGSIGETPKGTSIVRAVDIGHGATSPVYVSLSPDKSTO

CPn_0112 144743 145093
gatB-(Pet112) Glu tRNA Gln Amidotransferase (B Subunit)
DSDFGVNMKKNTHPEYRQLVFDSDSTGYQSEKTEVFEGKEYPVCYVSVSSS
SHPEFTGSKFVDAEGRVDFELKRYSNVROPAPQPEEDALPAAGKKKKVVTKKKK

CPn_0113 145329 146405
pFRA-Peptide Chain Releasing Factor (RF-1)
ADGKKKVAEYIKPDAVLEIKISNPFSNSKEYASKEHSYLLLEKNADNYDKITLNLEKVL
HDDMKALAEIKPDEPMVMEIGENKNEVLEKINKTLESLLVPPDPDDDLNVDINMELRAGT
GGEFAALFQVGDVCMYMHLYASSGKKGYEVLASESDLGKGYEVMGISTGQVKRLLYQEA
QTHRAVRPVETETQGRVHTSAITIAVLPEPSEEDTINENKDKIDITFRSAGGAGQHVH
TDSAVRITHLPTGVVQVQDQDSRQHKNKRAMRIKLRPIDAEMQKHNHEASMSRAQV
VSSDSRFRIRTYNEVSNRVTRDHRIGLTLYNLDKMEGLDLPRTITAMVSHAYHOLLEHGN

CPn_0114 146371 147261
 CME-A/G specific methylase
 VMPTTSYNNMEIKKAIQBGTYLDVYGVPLSDCEALYILMDLLEVSSRAKFLDLVGISET
 MLMEYRKRLALQRCRLPTALYKSGVFLGLRLVDSRVLIPRTEITELLAEYIIVNLLS
 EQITFYEDICGGSGGCGCAIKGSCPHVEVVLSDVCPVAVENAVENAKSNGLDGLDGLS
 APYTRPADAFVCPNPGYSLFNEIHHIDPEVRCYEPWKALVGGSTGLEFYQRIARQELPKIVT
 STGVGVLEIGGSSGGESKNIFSKHGKIRHLDLSDGRDRIFLEMDGRDPVSSGAS

CPn_0115 142779 148622
 vfh-Signal Recognition Particle GTPase
 MINSLSKQSGSTTFFSLFVSRNNEENISFSEIRVRLALDADVNVHVVKFIISKVKEKIL
 EGEIKHWHSQGLQVQICRLCHEELVAFLSDGREEFFIQKTPSTIILLCGLQAGAKTTHAAKLA
 YVQIKKKAKKVLVLPDCLKRAVAADQLKLQVATQKAEFYQSQENKPINVVVYLAKAYAKE
 NGDHFVLDITAGRLNLIMELMELTAIQKVSQANERLFVNMVAGQDVLTAQVAFDQSLD
 LTGVLSIMDARAGAVFYSIKHVLGKPIKFEQCGGERTQDLRSPDQSMALRIKGMGTDI
 LNFVKEKREYISSEEDAEGLKVLTAATFYEDYQKQAFMRFLPKRLKLLGMPGFNNAKIP
 SKQETIEDKQMSQAEVAILIILSTPEERKELVELDMSRMKRIASCGGLTLGDVNVQFRKQMS
 SKSKFFKQSGKGMDEVRKLLSGGNGNWR

CPn_0116 148592 148972
rs16-S16 Ribosomal Protein
EKNVRRKSVALKIRLRQGGRRNHVVYRLVLADVESPRDGKYIELLGWYDPHSSINYQLKS
IRFIYWLGERGAQLSSKAEALVKQAGPGVYSALLSKQEARKLVLVVRKKRRAYRQRIRSTQREE
AKADTK

CPN_0117 148983 150071
 crmd-cRNA (guanine N-1)-Methyltransferase
 GTGKMDIILSLSPFGYFDGLPQTSILGRAIKQRLLDVLQTLNLRDGLGKWKQVDDTPFSGGG
 LLMLAEPPVTSALSRVKNKSNKSVLYLSPGQALLTAEKRSLEAAASHILLCGHYGIDERA
 ESEVDEEDISIGDVLVITNGCIAALVLIDAVSRFIPGVGLNGESAERDSLNGLELGGPQYT
 PRPFEGFGEVPEVLLQDHNKASQWRLESGVLTETTYRRPDLVLNLYLKRSIDHKFDEET
 NPNRHFCKDKISVLVLEKLLKRAKNFYCKVGLGDAMSCENKFCPLPHEGKTIFWLREVQAE
 KSNKIVTSLSLDCACEEDFYLLRRRWELFGGKLLLEKQADEHAVWALAQDLGDHAWIFSWH
 KMK

CPN_0118 150075 150464
L19-L19 Ribosomal Protein
KCNFRWYIMVNLLEKELEQECRNDLPEFHVGDTRIATKISEGGKERVQVFQGTVMARR
GGGSGETVSLHRVAYGEGMEKSFLLNSPRIVSIEIVKRGKVARARLLYLRGKTGKAARKV
FVFGPRSSKK

CPN_0119 150520 151164
rnnB-Ribonuclease HII
MNTSISEIQRFSLMAIEFEKELVSEDFSVVAGIEAGRGPLAGPVVASACILPKGVKVPFG
VNDCKKLSPKQRAQVRDALNQDPEVCFGICVISVERIDQVNI LEATKEAMLQAISSLPI S
PDLILVDGLYLPDIPCKKI IQGDAKSASIAAASILAKEHRDDMLQLHRLPYGYGDFRH
PEYGTDLHVEAIRRYSPIHQKGSFSPKQMCIV

Frn_9120 151125 151778
frn-2MP Kinase
ELFQFNKANVYQCNKILVDFPSFDHKCKPKLFTICAPAGVCKTTLVRMLGEFSSAF
ETICVTYTRYPRCEIVNKKYHYFVSHFFORILLDQALLENWVFLNCEYGTSMLEIERIW
LQGHAVAVTDIQGLFLRPMPSVIFATPPQEELERRLASRGCEGQORKERLEHSL
ELAAANGDFVVIINDDLGKYYRVLEKLTACEIHNNIL

199_0121 151769 152068
 T031 hypophosphatidyl protein
 HIIIMIKKDEFTTEKNLKLIVPEFLVNFYIAIKQAKIFIAKGDAVPEENVAIETLVLLDREGI
 PEFTFELVVTATITVERKKIILINIPKKDIIAYTWGDKV

•

QKYMQILLDAPVSLLYGAFKGFKNCTULNITELNLANSTKAAEEEAERYVEEKGRGFETY
WEEAKQRLEAIAAELDDLNRQETLLEQETPLANKIETISDLNLRKVSVEKAALEETIQ
GIEQYAEZMGIEDLELKQKFEDLQKKLEALERLLQIGRRIDSSVDKQKELLGLGREE
AA

[illegible]

CPn_0152 195274 194318
 C149 hypothetical protein
 LKMKRVAVLVCSLFSVAGASAPRVVPGFQIPEDLVQIKTEVCKPEQVCLAVTIKCD
 DHNLKGLVHLNPTPTPEGQFTPTVLPHFGFRQTKFGLGTAYRKLGRKFAAVGIALRVDM
 AGGCDSEGVAEVEVETPLVLRDAQITLVEQHPDLNAYRLLSGFSGLGCHIAFELAKIYN
 PRDLNLKSVNAPADGGLLKLELVNFSKHGEGDLSVQKGDGFGFPPPIVCSQDVL
 LRITQDHWANSLPTKPKYILHQQIDDTLVSRQTQTLFKNTAPGRMTFISYPNTGHNLAT
 APRLDMILDOIVSHFRTL

Cp_n_0153 195430 197892

leuS-Leucyl tRNA Synthetase

NMRYDPNLIKWKQWQFKWHERSPQANEDDEKVKYVYVLDMPFYPSPGAGLHVGHILGYTATD
IVARYKRARGFSVLHPHMGWDSFGLPAEQVAIARTGTHPKVTVTKQNIANFKKLASMGFSYD
EGREFAITSDPDYVYHMQCKLFLFYLDGOLAYMADMAVNVPELGTVLSNNEEVNGFSIBGG
YVPERKMLRWQLIKITAYADKLEGLDLDPENWVKQLKNWIKSGEGALVTFTLHTQGS
LEAFTLRDLTDLSVFLVIAPEHPDLDSLVISEEQREDEVYAVQESLRKSERDRISSVTKT
DGVGFTGNVAKHPIGNNLLPWISIDVYVLYGGTGMVGVPAHDEREDREFAMFLSPRIHEVI
DDNGVCITHSNYNDRCFLAGLSQEAADYVINYLRLEKCAKTRMYRLRWLFSRQRYWEG
IPTIHFEDGTHRPLEDDELPLLPPNIDYVREGFGQGPLAKAQDWHIYDKETGRPGRCRE
TYTPMOWAGSCWYVLRFCDAHNSQLPWSKESEYWMVPVDLYIGCAEHAHVHLLYSRFWHR
VYFDAGLVSTPEPKFLLINOGLVASSRYIPGKYGVSIEDVREENGWTSTCGEISVEVRQ
EKMSKSLKNGVDQVLTIEYGADLRMYAMSPGLDKNKLNSNEGVGDCRFLNRFYDLV
TSEVEDITDRDGLVLAHLKLVFRIITHEKMSLNTIPSSFMFELNDDFKSLPYVSKRALSM
AVRVLPEIAPHISIELWVILGNPQGIIOAAQWQIDESYLAQVTVTFVQVNGKLGRLREV
AKEAPKEVYSLRSVAVKYLENAQIPKRIEIVYPNKLVNFVL

Cpr_0154 197874 199202
gseA-PKDO Transferrase
TSEFCPMLRLGVHRIKCFYVDVVLVCAFFLKKLYKMLVYGKYKSLAVRFLGKKPHV
RGGPGLVWHFCAISGVEVRLLLPVEKFEFFPGWRCLVTSCETLGVOVASVFGPMGATV
SFLPLDFDSIIKISVAKLRSLPVFSGGDWLPIEAKRIGATTGVINGRISIDSKRF
KILRLKGNYSFVPDGLQDEVEQKORFLSLGIEPHKLGITGNIKTVYAAQTLHLERET
WRDRLRPLTDSKLVGLGSMHSGDAGKVLVVKOLIKEGVSVLVWPHRPHKGTMDVEESLHR
LHLPYGLSMRSGANFSYFVPMVDEITGLKLVVAGDLAFVGGTFDPKIGHNNLELPQCE
VPLIFGPHITSOSELAQRLLSLGAGCLLDEIETIDTVSFLNNQVEVRAEYVQKGKVFVK
AETASGFDRTWRALKSYIPLYKNS

CPr_0155 199697 199488
No robust homolog present in Genebank/EMBL as of 11/7/98
NSLSFGVPFLEKLKISLIPIEEMRHFLFMKTHNSSNGFSNQEKGIRTYFKSDLLGYEDL
YFTRENTNPN

CPn_0156 200147 199770
No robust homolog present in Genbank/EMBL as of 11/7/98
LGKQKLLARHMMHNIVLSEEPGRSAFLGRTAFFPNKYPIAQGGVGIPSTIGNLFTIWYC
FYFYRAATPQSDHPDGCGFILLERLKLGAAGFFYCDLRESNTTGTFTLFEGSNKGVLKNH
LGRIDE

CPn_0157 200753 200298
No robust homolog present in Genbank/EMBL as of 11/7/98
FSFVITYKEALNNIYQFSGASPNWQASLMAQLNSYFCLGGETVTRIISLRPSGLILAKKE
KAVVSTAEKILKILSFILFPLVLIALAIRYLLYNNKFNKDLRAVFIPITEITKAEELIIA
KNPALVKEAALTYSPLFYSLPLPKKYQIMKTVEL

CpN_0158 201463 200894
No robust homolog present in Genbank/EMBL as of 11/7/98
PKPITLSINIDLLLEDLDITDSIPWPKLYLSEDFPAYEPESKAIITVALEKKNNGEEF
CLESKKILARYLLEQLFKLETGLNFPSTSIDGGRESFLIEFSHETKKKPTWAFIFYFYYH
SNGPKLEKDFKQAGCEVHNRLNLNLGLKTRPQAGAQNDRNGGPGYGI GLFVLWEENYGSV
LKDGHGFKIDKN

CPr_0159 201811 201467
No robust homology present in Genbank/EMBL as of 11/7/98
CCPFGSTATRIFSMTPSGFSLATEEK/QVSTAEKVIKILALIFFPIILALAIRYFLHRK
FDRKCFVIPQDTPKELELILANPOLVEKAAREVHPGFALPTKYOSMYIOTSKG

CPkA_0160 203794 202127
 tvELLSS-LNtXSE-6-P Phosphotransterase
 TVELLSSLNtXSEYFEIQLRYPREILTLETLSRKIHQETSSPPSPPELQKHIENLCRIPE
 VSIYKDEETSSPKLIGVLGGGQAPGHHVVLGLFDALRVNPKTLRFLGPKIGKPLGLSTR
 GYKGLDDISVIVDYNNMGMLDSSSPCEKITEEOKNTLNTVKQLDGLLIGGNNS
 TOTAMLAIEFLAHNCKTSVIGVPKITDGLKNCWIELSGLFHTSCRTYKGNENIGMLAKDAL
 SAKYVHHIRLMQGSQSYTLECGLOTLNPLALISLIEATRIKLSKQLSEQLGLGLVRRY
 KGGKNYSTVLIEPGLIEHTPTPKLIELNLVLANGDGSICKLSKLSPELTKTFHLFPK
 DLANQLLGLHGNVRRVSKATEELAAVNAKKCEIKIPMHMTESVHHFGEYEARGGT
 SNFNDNYGLALGISLVLPRPKPTVMITNNLAKYVTEWQKATPIYKMHMLHRCGTE
 TPTVLTSDVDPKSPAVQHLQQGDSGLVDEPLYHFGPLQYFYKKEELIDQRLPLTLENQIT
 HSEFQALYSTSGKRLS

CPN_0161 201658 203704
(predicted acyltransferase family)
HRSSTRRKQEPLLTAVVLEKKHEQPMSTLTLLNNPTTFGLIITPTTIYNPYFIVILLGG
LALDKTSGSRRHRLAQLGELTRGLAALPVLVLLIGDCEYELMDFEIPNQRKQETIIEY
HGLILDQERLAIPGSLVPLALQVLPFNKIKALAWAATPTIHLMAAFAQNAPEVI

TM:QKGAITYACMTLNPDFYTOFLKIDIVKELMPSARNLPILYMOGEQDILLVINHRTL
PTEAFANQDKPITILTYPDVDAHAFPAESSALDQLTQWLKRELTSGE

CpN_0162 205870 204803
No robust homolog present in Genebank/EMBL as of 11/7/98
FVYTLNYSQSPFRIMKLYSISSDVTPWIFQLMSKVDYSYFLQGNRIKVVSVIMQEPNLI
IKGVENRISTIVKILKLSLFIPLILIALALHYFLHAKYANHLVSKILERAPOVYPI
KVRGDTASHYKLTTLVPVQSNKQAMQSNPVEEAALRTTKPSFFCVPAKYRQIIISSH
LRFKILHIALYINLDAWITLITIMD: *ADKPLQNMJLPTTYINTVQKP
LRFKILHIALYINLDAWITLITIMD: *ADKPLQNMJLPTTYINTVQKP
PLEDRGGGFEILEQLQELGVRFPICSPQGDNPFPQGGQIRIYWEDSYQPNKEV

CpN_0163 205831 206394
No robust homolog present in Genebank/EMBL as of 11/7/98
FEKAIYVICIKKQIKICISIIHTPATPLCTEGEIFFPGPVDSAIQNDLERLLTVKKRPD
IIREYLRAGGSLVTTYKPGQRLRSPEQLRVLDDLVQSYPNHLHAIELDCGAIPODLIGA
TYIITFADGSTYILSLRSYQANSPSDDTWIGFWGSIIDDPQAVISFLKDHGFALPSTLAQ
DPLLCNTNK

CpN_0164 206444 206998
No robust homolog present in Genebank/EMBL as of 11/7/98
LCFKCIYIKIIFSLKQMLTRSTIESSDSLCSRSFSQKLSVQTLKNCESRLMKITSLVI
AFLTLVGGALIALAGGGVLSFPLGLILGSLVLFSSYILVSCCKFTLKEMTMCVSYS
KINIWFEKQRNKDIEALENPDLFGENKRVGNRSARNQLEMILHETDGIILKRYMGAK
MYFYL

CpN_0165 206983 207582
No robust homolog present in Genebank/EMBL as of 11/7/98
NVLLFMNWPKTIDHVDPESEIDIRKVVSCYKLIKCEQPEFRSLISELGVIRCGRLRLK
RSKYQEQAARTVDEDAFLCLTRSYQDGYLTPLAGPRDLINHYIHLRRENPKHFSP
KHCYCYARLAFNESVCYRLEFDIERLTKMYVEGDYSKEQEKNLQAILSFVKTLEDEKDF
LIEHKDITDLIRGFTDVFCT

CpN_0166 207594 207962
No robust homolog present in Genebank/EMBL as of 11/7/98
NCLKQYNKSDSISMESINRSIHLEASTPFFIKLNLCESLRVKITSILVLSLALVAGAVT
LVLVFAVILPLLVILEIILITVLVLLFCLVLEPYLIEKPSKIKELPKVDELSSVETD
STL

CpN_0167 208309 207977
No robust homolog present in Genebank/EMBL as of 11/7/98
NLWSHFPGRGFMLPFCPTILLAKPLNSENVGLERLAATVDSYDFLGGQSIVFLSKQDQ
ITVEELSAKDRKFKPGSMNCTLYTEDPILPAHNSFSCSDIQMRTPISPIH

CpN_0168 208716 208417
No robust homolog present in Genebank/EMBL as of 11/7/98
SYINLRRENPEHFFNPGHPCYARLAFNESVYRKLFTALQMYGAGDYEQNEDN
LKSILSFVQILDEKDFDPLKADTTFIGRGADIFCS

CpN_0169 209537 208710
No robust homolog present in Genebank/EMBL as of 11/7/98
SPHTEFTIGENMMQVGESECSQPLVMELNTQPLRNLCESLRVKITSFVIALALVGGITL
TALAGAGILSFLFWLVLVGLVLCALFLPSYKFCPIKELGVYNTDSQIHQWQQRN
KDEKATENPELGENRAEDNNRSARQVETLRDCDGNVLKKIYERNLVLVLFMMNWPK
TMDDDVPYSEDSIRTVISCYKLIKAKCEPFRSLISELLRAMQSGGLLSRCSRYQERAKT
VSHKDAFLFCPTHSYRDLTPLRAGPRYINRAI

CpN_0170 211098 210025
No robust homolog present in Genebank/EMBL as of 11/7/98
NVRKNHIIIRGEKYNCTVIAFVLSMSYDTLFLNLEKEDSVHKICNEIFALVPLNTIAC
EALKNLKPADIHVHLPGTITPQLAWILGVKNGFLKWSYNTNHLRLSPKNPHKQYSNI
FRNFQDICHEKDPDLVLYQNYLNDPNSFDRVMATVQGHFPPGGQIONEDLLIFNNY
LQGLDDTIVYEQONIRLAHLVPSLPEKHARKMFQIYLVRASQTFSKHGITLRLFLNC
FNKEFAPQINTQEPAQEAVQWLQEVDSFPGCLFVQISAGSESAPACPKRLASGYRNAY
DSGCECAHAGEGIEIETRTIFSSAKVNPGLIEIETRTVSSLRKQKPPSSLPIRVTCQLG

CpN_0171 212444 211149
*guaA-GMP Synthase
IKLKQSARRHLNTIFILDFGSGYTYVLAQVQRKLKVYCEVLFWNISVQCLKERAPLGIIL
SGGHSYVENKAPHLDPETIKLIPILAIACYGMQLMARDFGTVSPGVGEFYTPHLYP
CELKFHIVDCESLDEIRMSHRDHVTIIPGFWNIASSTQCSISGIENTKQRLYQLQPHW
VYSDSTPTGNKILETFVQIEICAPLWNPYIQQDLVSKIQDTVIEVDFEVAQSLDVQWL
AOGTIYSDVIESRSSGHASEVSKSHNVGGLPNLKLKLVLEPLRYLFKDEVRILGEALG
SSYLLDRHFPFGPLTIRVIGEILPEYLAIRLRAIDLIFIEELRAKLYDKIKQAFALFLP
IKSVSVKGCDSRYGYTIALRAVESTDFMTGRWAYLPCVDLSSCSSRIINEIPEVSRVYD
ISDKPPATIEWE

CpN_0172 213237 212440
*impD-Inosine 5'-monophosphate dehydrogenase (C00H-terminal
region only)
APIGAAIGIGPLGISRAHHLVEAGANVLVIDTAHAHSGVQFTVLEIKSQFPQISLVGN
LVTAEEAAVSLAEIGVDVAVKVGIGPSICTTRIVSGVQYPOITAITNVAKALNSAVTVIA
DGRIRYSGDVVAKAAGADCVMGLSLLAGTDEAPGDLVSIKELFKRYRQMGSKAMQKQ
SADRYFGTQQQKLVPGQVEGLVAYKGVSHVDVLYQILGIRSGMGVGAETLKDCLKTKAS
FVRITESGRAESHIHNIYKQVPTLNY

CpN_0173 214041 213715
No robust homolog present in Genebank/EMBL as of 11/7/98
TIFDLIYKIDSYKHQGGFMDFSVPPDRFVSTSPSPEDIAKTLVSNCHYCSRLCIF
LGLLSIIICFSVYCTSGETASLVPGILSLILVLLIIECRNRECCRIS

CpN_0174 214215 214724
No robust homolog present in Genebank/EMBL as of 11/7/98
FIFWTFEIKVILGIMMTTICNSPSALNPESLIPPTLVGQXTQTSLAYTIPAQGRS
TLRIILDIILILGLATICTFIVIFLNLGLLTSTIISGKLIIGVLLFLIMGLYFM
LISLDYILVGLQLQELQAEEREEYIQTIEALRGAPRAECPTESPSTWL

CpN_0175 214854 215275
No robust homolog present in Genebank/EMBL as of 11/7/98
LILACPOFLLRKPDMEQNCYIQDTTVLYALNSLCPRLSDTHRLQKQPLEAENALGE
FIFWTFEIKVILGIMMTTICNSPSALNPESLIPPTLVGQXTQTSLAYTIPAQGRS

FLTDGMSPELLSEVKEALFP

CpN_0176 215271 216518
CT153 hypothetical protein
NDDPMDESDGEEAKSDASFSYSEYFVKSSSTRESKNTVTHSTASRTLYLRLQDCSYDP
RALKVDDEFRVWVEKRLDAKNPDSNAFVKEVGHYVASVTVGGIGFOVLKMSYLOVEEL
EKEKISISVAASSLLSKSTNATEKGYSSQESSAQTVFLGGTVLPDLQDKDLDFKW
SEGINPEIPLAISVSSITDLIIPELFSPSEDAQVLSQKKSALQGVILNLYLESHKPKKEGP
LRLKNTTGDILRIGDEIFLRKINTSUKRLATTSMJGALPCTTTTSSSVFIITV

CpN_0177 217513 216608
No robust homolog present in Genebank/EMBL as of 11/7/98
DKREQTKSKFIFLISEESMKQMSLIFSSVCLGLGLCSLSSCNQKPSWNYHTSTSEEF
VHGNKSVSQLPHYPSAFRTTQIFSEEHNDPYYVAKTDEESRKIWRIEHKNLKIKGSIPI
STYGLSMHPKSAALTLTKYRPHPIWINGYERSNITDTGKYLKNGSRRTSHDGPKNRAVL
NLIKSSGRRCNAIGLEEMTEEDFVIARRREGVYSLYPVEVCSYPCQGNPFVIAYAWIADESA
CSKEVLVPGKYSVLWESVSSDLSLNAFGDSFAEDYLRSTFLANGTSLCVHESYKVPVP
QP

CpN_0178 218052 217789
No robust homolog present in Genebank/EMBL as of 11/7/98
VKEYLDLVLQVRNVERDPQTKRCHTVSQKFGESIDAKTTTGQLFHIAGKTEPGHKLCLG
ESILKQLLALGIITGYENREREVWYLD

CpN_0179 218550 218056
No robust homolog present in Genebank/EMBL as of 11/7/98
PKIWDTHFETRIEATSVPKFNRLRKSFHKSGRSPRSKACVANFFNFTLQAGRSIIPG
KKAILLNVDAKTPNYSCIFESIGFFNEQDLEAQNQQAAVRLKILKVPHHFKGLIAK
LPRSALKDRKFMSSLIPTKLSYALDLSAPMHLEKPNLSYEKLD

CpN_0180 218963 218355
No robust homolog present in Genebank/EMBL as of 11/7/98
TSLHILDCYKYPVFIQNTVASSETYPSQILHQAQREVVDAYFNQADCPARANQILEAKKI
CLLDVYHTNHYSVFTFCVDNYPNLRTFVSSKNMNGLSNPLDNVLEAMVRRTHARNL
LAACKIRNIENVRVGLDLRSGILISKLELKPQFQSLTEDFVNHSNTQEEARVHQKHLV
LISLILLCKQAVLESFQEKXSS

CpN_0181 219175 218777
No robust homolog present in Genebank/EMBL as of 11/7/98
VHELFFKIDGVYFFFKFMKLFPNNYSLNSHHEKSSLEKAVQALDSYFYWGDDTDLVAR
DDISREYCVRLYIRFWIVSISQSLSRIPWRLKRIILRYCTLRGKYVMPILIKRIAILL
GLIRFSRLRSVY

CpN_0182 220704 219334
accB-Biotin Carboxylase
RCIMMKVLIANRGEIAVRIIRACHDLGLSTVAVYSLADQELHVLLEADEAICIGEPQAAK
SYLKISNILAACEITGADAVHPGYGLFSENANFASICESGLTFIGPSSSESIAMKDKIA
AKSLAKKIKCPVIGPSEGIIEDESEGLKIAEKIGFPIVKAAGGGGRIRIVKEKDEFY
RAPSAAAEAEAGFNNPNVYIEKFIENTPHRIEIQVIGDTHGNYVHLGERDCTIQRNRQKL
IETPSPILNAEIRVKGKAVDLARSAGYFSVGTVEFLDKDKKFFYFMEENRITQVEHT
ITEEVGTDLVKEQIHWAMGNKLPWKQKNIPEFSGHIQCRINAEDPTNNSPGRGLDY
LPPAGPSIRVDGACYSGYAPPPYDSMAKVIKAGKNREAEIAIMKRALKEFHIGGVQST
LPPHQFMDLNPKFLESNDYINIDNLLAQSSFFKEF

CpN_0183 221207 220695
accB-Biotin Carboxyl Carrier Protein
RRLGMDLQIEKLMIAMGRNMRFAIKREGLELELDRDREGNRQEPVFDYSLRFGSFS
QERPITVDPKDITIKETTENSESTTTSSGDFISSPLVGTFFYSGPAPDSFVKPGDIV
SEDTIVCIEMKVMNEVKAGMSGRVLEVLITNGDPVQFGSKLFIKADAS

CpN_0184 221814 221221
efp-Elongation Factor P
QWKIKFCCEEKIMVLSQSVGMFISTKDGKLYKVTSVKSVAGPKGESIKVLAQAADSD
VIERNFKAQTEVKEAQFETRTLEYLYLEDESILYFLDLGNYEKLPIQEIIMKDNFLPKA
IGTVSAMVYDNVSVLEPHFLELMVSKTDFPGDLSLSGGVKKALLETGEIEMVMPFVE
IGDVIKIDTRTCYIQRV

CpN_0185 222457 221765
rpe/araD-Ribulose-P Epimerase
AEVVKQESVVLVGSIMGADLTCLGVEAKKLEQAGSDFIHIDIMDGHFVPLNTFPGGIIAA
INRSTDLFLEVHAMIYNPFIEISFVRSAGADRIVVFEASEDIKELLSYIKKCGVQAGLA
FSPDSTIEFLPSFLPCDVVLMVYVPGFTQGSFLPNTIEKIAFARHAIKTFLVGLKDSCLI
EVDGGIDQQSAPLCRDAGADILVTASYLFEADSLAMEDKILLRGENYGVK

CpN_0186 222878 224068
*similarity to Cps IncA
PIKDKILMSPPVMTSPAPNIPAPPTTPIPTTKPRSSFIEKVIVAKYILFAIAATSG
ALGTILGLSGALTPOIGIALVIFVSMVLLGLILKDSISGGEERLREEVSRFTSENR
LTVITTTLETSEVLDKAAKDQLTLEIAFRNENGNLKTAEADLEEQVSKLSEGLEALERI
NQLIQANAGDAQEISSELKILSGWDSKVVEQINTSIQALKVLLQEQWVQEAQTHVKAMO
EQIQALQAEILGMHNSQSTALQKSVENLLVQDQALTRVVGELLESENKLSQACALRQIEI
LAKQVHESLQORIDAMLAQEQNLAEQVTALEKMKQEAQKAESFACVRDRDTRFORRETP
PTPVVEGDESQEEBQTPPVSPQSSPVDRTAGDQ

CpN_0187 224218 225045
predicted methylase
VFLTYRTLPMHSHKFLSRKKNSSHKKEETQWDCIASSYKIVQDKGHYHRETILPOLL
SLTSLGSSVLDDICGGQGLFEPALPKEDRYLGLIDISPLIALAKMKNVNSHQFKVADLS
KRLFEVPTLFSHAVAISLQNMFPQGAIRNTATLELLELQGFVLVIANHPRIPRASSW
HYDEIKKAIISRIIDRYLSFMKPIIMAHYQKQDPTLQPHFLPQWPKELSSHGFLVSOL
BEWTGKTCTGKRAKENLRKEPFLIMCTIKIK

CpN_0188 225090 226406
CT152 hypothetical protein
KTLNCKIMFRKLFPFHKKTGQFORLKNLILQAIICQILVILHNLHAKKEACVLYYGLL
TVVILVFLPFLDLQHLPTNLAWYEWLIIKFDLFFKPIVAVIAAZHATERNIGLVAGCF
VWYAGTLMILSLDGLNKLIPPTWPLILYPLVYITPLVAGMIFITVQEWIYITQ
IMPQYAKLGLSHSMALYFCRFVYVLLLYLALYFZAPLPPVAKQKHALITPLIG
WVLYPQKAFGLQVIFPNYSFTGALVALPFLILLIITYMILYFKYALFTIQNRCT

VROWEKIAAAKFCREQGIPIYFGICLQWVUVEYARNVNLQDANGSEMDPNTPHPVIV
VMEQDQFLVATGOTMRLGAYPCLLKPGSKAHKAYNESSLIQERHRRHYEVNPDYIQSLD
HGLRIVGTCPPQGLCEIIEVGDHPWMTGVQFHPEFVSKLISPHPLFIAPIEAALVYSKDA
SHV

CpN_0237 273741 274214

yggF Family
GILRMDAMSKPSSCKAYLGIDYQKKRIGLAYAAEPLLLTLPIGNIAGKNLKSAAELHK
GILRMDAMSKPSSCKAYLGIDYQKKRIGLAYAAEPLLLTLPIGNIAGKNLKSAAELHK
GILRMDAMSKPSSCKAYLGIDYQKKRIGLAYAAEPLLLTLPIGNIAGKNLKSAAELHK

CpN_0238 274210 275838

zwf-Glucose-6-P Dehydrogenase
PCNHQKLRDPNFRNFFLVIPASAGTKKEIKMTNVVQETIGGLNSPRTCPCLIVIPGAT
GDLTARKLLPALYHLTKEGRLSDQFVGVGFARREKSNELFRQEMQKQVIFQSPSELDIKV
WEDFQORLFYHRSEFDNNMGYTSKDSLEDLQKTYGRGNRLFYLSTPPQYFSRIETENL
KHKLKYQDQGGKQWRSVIEKPFGRDLDSAKQLOQCINENLNSVYHIDHYLGKETVQ
NLTTRFANTIFESCWNSQYIDHVOISLSETIIGISGRGNFFEKSGMLRDMVQNHMQLLC
LLTMEPTTFDADEIRKEKIKILQRIISPFSESSIVRGQVGPOTVQGVSVLGYREEENV
KDSRVETVYALKTVINNRPLWGVPPYLRAGKRLAKKSTDISIIFKSPYNLFAAEESCR
PIENDLLIIRIQPDEGVALKFNCKVPGTNNIVRVKMDFRYDSYFQTTTPEAYERLLCOC
IIGDRTLTGGDEVMASWKLFTVLEEDWDQSSPSFPNYAGSSGPKADALIERDGRSW
RPL

CpN_0239 275863 276672

devB-Glucose-6-P Dehydrogenase (DevB family)
KSISMTNIGIETMATLINFNKLLTKQPSLFDLASKDWIASANQAIKQAGFYVAL
SGGKTPLEIYKIDVINKDLIPSKIFLFGWDERLAPITSSSNYQAMSLRLDNLIPDE
QIFRMETENPDGAKKYQELIENKIPDASFMLILGLEDGHTLSLFSNTSALEEDNVV
FNSVPHLETERTMLTFPCVHKGKHVVVYQGENKPKILSVFFSEGEEKLYPIERVGRD
RSLPVLITSPESYDIADFDNISSYIKMDIL

CpN_0240 277861 276698

No robust homolog present in Genbank/EMBL as of 11/7/98
LVYFMVSPSSSEVVKANSVVRNFCYFLENKVFSPSESTEVMFSEIMKGRVPDIESLFD
RPTDMMTGFKAAQNLGNLNSFGILLIMCFSQCKSCQTEPEKETSALVIGATLLFFVVALI
LGPTLGLAVYCAKYVYTLGKMIYSLNKAQKVLKRAQVNHRAAGVATIRSEEAVKAC
KLYKSAMTIGSLVSLIASLALIALTAGIVLVLFVFAAGAAPIVTAAMGCCAAGGALLI
SLGLWIAIVRKAHQEACVGHLMVNLVLTAVSEALLHDPSPHTNALARDLFLTDCLSH
YGHLSNEEVAQLVQGGAPGGGSRFSQYGGSSDYQNRGGNGNGFGSHFGGGGGFAGSH
FGAGYPTAPTMPSAPPPPPPPAYDTIYG

CpN_0241 279372 278203

No robust homolog present in Genbank/EMBL as of 11/7/98
IFIKVFKSAMISLSSSHEASIASNTQVRDVLVSLAMDEFVHNTEILPIKVLARGTLSS
TALIDDLKDVETEGEHHFQVYNSISLKMICYQFEFKIFGICCPPLLVTDSDDPCGA
LITGFAAVLFTVLAIVFGPTLGLICYSAYKIQYLTTKISSLRTHTEVINSVQKSDPFI
HREAAVAAAASOSTIKACKVFRQSTLIFVGLIITISLAAIVGLVLFALFDLPAGAP
VMTAAMIGCCAAGGTGILLSVIGFLASVYVQKSGQGVHMHMTALLRCIVSNTIIQMPY
LPIETGTTKVLVTQSIIRRYQFFSDDEYRDIIESEVPLNRQTTTPPSYETLPHBEGSDGSSN
VIPRESPPAYSTIDSSNSPFPSSSSPPPYR

CpN_0242 279975 279487

No robust homolog present in Genbank/EMBL as of 11/7/98
KSLKYCSLYQFSQKPTVILMACSIFFRMSQGDYDDEPLSKKTACLVDVMTLYPVTAVCA
VSVVLLILKLVFLLSFPFKLCSSASSALPGRVSRVLSGHFKCLYGGGLPYLLACLILVPV
IGTAHGHFIISHRTSEDARLSSAIVFMQAPILQLAGMSGLIKP

CpN_0243 280609 280133

No robust homolog present in Genbank/EMBL as of 11/7/98
INENLVLLKPVKGRIIMACSIGYHLCNANEPDRFVASKVLADILLYPFAVICAVV
FAVLVWVLLFLAIFLNTCIAACKSRPLPSCKENFQCLFGPKDKPGPSDWLGLCLVLI
IIGLFIYSTIITVQSDTNRLRYFIISAPYQVGSTAIINW

CpN_0244 280906 281556

adk-Adenylate Kinase
GAPSPTKGSVFIIMGPPGSGKGTQSQYLANRIGLPHISTGDLRLAIIREGTNGLKAKAY
LDKGAQVPSDFVWEILKEKLOQACSKGCIIDGFPRTLQDQHLDSFLMDVHSNTYVIFL
EISEDEILKRVCSRFLCPSCSRIYNTSQGTBECDCVPLIRSSDDTPEIKERLTKYQE
RTAPVIAAYDSGLKLCRVSENKEDLVFEDILKCIYK

CpN_0245 281627 282499

ydH-Polysaccharide Hydrolase-Invasin Repeat Family
TCQKEIMKHLYLSFSPSADFFSKQGAETQVLFGERVLVKGSTCYAYSQLFHNELLWKPY
GHSFRSTLVPCTEPHIHPNVSVVSDAFDLPGWIPLPFGTLHLVNSQNTVIFPKDILNH
MNTIWGSGTQPCDPRHLRLNLYNFAELLIKDADLLNFPYVWGGRSVHSELEKPGVDCS
GFINILYQAGQYVNRNAAQADCHWISSFENLPSGGLIFLYPKEEKRISHVMLKQDSS
TLTHASGGQKVEYFIEQDQKFLDSTLYFFRNQGRGAFFGIIPKRKAFL

CpN_0246 282955 282551

rs9-S9 Ribosomal Protein
VVAKSTIQESVATGRRKQAVSSVRLRPGSGKIDVNGKSPEDYFPLEIQRITLSPKLIK
EDQSQVDLIRVSGGQIQGQVIATRLGLARALLKENEENRQDLKSCGFLTRDPKRRKK
YGHKARKSPQFSKR

CpN_0247 283430 282969

r13-L13 Ribosomal Protein
DGYIIMEKKRDKTKITIVKSSSETTKSWYVDAAGKTLGLPSEVAKILRCKHKVYTPHVA
MGDXIVINAEEKVRLTGAKKGQKIYRYTYGYSIGMPEIFPENMMARKPNYIEHAIGKMM
PKRLGKQKQLKRLIVKQDSYETTESQKPIILDT

CpN_0248 284451 283650

ydcU-P-Loop ABC Transporter ATPase
KCHVYGLFERYVYATIPSPFPAKCKRKNAQLPHFYGLAMILLIEAKNLKTIQOQ
QNTILITPAVGLLHAYDTIITGAKNMTTLVLLHLLCTLOVPSGTLRFPPDKLQNDLA
NFKKHIIDVFPONFYLLDDTVLKNVIMFALIAKRIHFKGKPVYTRALELLDLYNLEDKY
KPKGKIKKZKFORVAIARALINDIALLADRPQHLDECTEIOHNLLLQASALCGIL
IVTHNHKLAHPCREGLNCKLPHIN

CpN_0249 285850 284133

CT141 hypothetical protein

LEVMKFEFSVALKYIIPGRGLYSAIVSLFTVGIISLVWVLSIVFIIVHGLQRWIEDL
SQLHSPITILPSDITYSYTYQIDKHSSLSN/TTKTLEKGIASPOVDYDPSDYLLPET
FPLKDCDLQOQKQDPVKMTLESGLPYLQSQHGKVIIEFEQGVYLDIKTSLKLOKQPPRNL
THFLTYPKSLKSYEDKVLVYDETDTYSAELNPFNRSPSGNQDFFHLEELYRGASIIIPST
YKDSGYKVGDTGVFSTYSENEKETQYTVHVIGFYNPGLSPLOQRTVFIIDPLARSISQ
SEGLGMSNGFHLFPNTKRIVFVKQIENILTSGLGVDDVWEISSLHDYDFQPILODQOS
DQVLFVLCILILIVACSNIVTMSMLLVNNPKKEIGILKAMGTSRSSLKIIIFACCGAFSG
ACGVVIGTIPAIITLKNQFIVKALNYLQGRETFNTAFFGQNLPSNVHPQAIYFLGLGLT
IIVNTHIAPYVNMHTFVYV

CpN_0250 285862 283302

r133-L33 Ribosomal Protein
KDDSMASKNREIILKSSSESSDMYVTKNKRKKTGRLELKKYDRKLRHVFKEAR

CpN_0251 286036 287559

*conserved hypothetical protein
SPDSCPLWMSPPFKIVNRLCYISFQKESRTLPPIIREPMTTKSLGSFNSVSKNKIHF
ISLGGCSNLNDEVMGLILLKAGYESTNEIEDADYLILNTAFKLSARDEAKDYLDHLID
VKENAKIIVTGCMSTNNHDELKPMWMSHIIHYLLGSGDVENILSALESRESGEKISAKSYI
EMGEVPRQLSTPKHYAYLVAEGCRKRCAFCIIPSIKGLRSKPLDQILKEFRILVNSV
KEIILIAQDLGDYVGLDSTDRSSQLESLLHELLKEPGDYLRMLYLYPDEVSDGIIIDLMQ
SNPKLLPVVDIPLQHINDRILKQMRRTTSREQLIGLFEKLRKVPQVYIRSSVVGFPGE
TQEEFQELADIFGEGWDNLGIFLYSQEANTPAELPQIPEKVESRLKILSQIQKRVN
DKHOKLIGEKIEAVIDNYHPETNLLTARFYQAFEPDPIIVNEAKLVSHFGERCFIE
ITGTAGYDLVGRVVKSSQNALKTSA

CpN_0252 288112 287576

CT144 hypothetical protein (frame-shift with 0253?)
ATSTVCAWLILQTYQSHDDAASSFRACRFGRYLVGGVNPWNKFNQTSQTSTVNSAI
YIDSSQTMWMPRQASASIPRLFRISIFMTKHGDWINDGTGGELLVAYEANOPLFPDIR
IELAMSTCSQSYRARPQWLCTSTYVLRPGYFVLENRSYFRVQSPFNSWNIATLFPVN

CpN_0253 288474 287950

CT144 hypothetical protein (frame-shift with 0253?)
FCGGRMLSSSIPPTQKITISIPTFVRNIESINLDEQKKTALTIGONIAENTQVLGNF
VDADGLICQNDLSVGGNINITPOTFNTMVFSGRVNLNSPFSYQDSLGKQDYANINT
BQOQYVPYGYKLTRVMMQRAALSGGHVSGSDIGWGESMYLGISSIKRQHKVQ

CpN_0254 289268 288459

CT143 hypothetical protein
IPMKTLGVKQNLFIQDQATLSVERNRIENNELTRDLKVLDTTSPCEFIVKGNVSAEGS
QLNATLTGSDNYSIKTDVSTQVPCNNISDPQASARDLTFSYRKTGQCAANLYTYYPGN
GYVVAPTNTIETHVAITSKSVSRNATPDFSRYSADIPEVVKLKQVGIYQVMTQLTRWSGQ
HDGNSATLILNFVSGNNKTLTCTSDTROGSSRDITSVAVTAIFSVELVSSPPYDYPWI
NLESTIWMNLSLSTCVIWFPPSPNFVEVD

CpN_0255 290183 289329

CT142 hypothetical protein
VTLKVMKNNINNECYFKLDSTVDGDLAANLKTFTDQAQGISSTETFSVQGNATFKDQ
VSATGLTSGTYVNLNAQNTSSQISIDFNKRLSNALPKEDCDPVANYRSPPEYFCS
KPLIGDFDNGSESYLPLTGSEYTLYQSRNVSIFRIGWKQSTRELTVGGNTAIQFLAA
GTIYVSFTVGRKWWNGWGGAIYINNGLGQVQCESTIYSGGYATIGTLGTSIYRASVD
VAPNPDNADSDRYRAGIYLSNGGSSAGIGNYSFLLYPPDRG

CpN_0256 291282 290398

CT144 hypothetical protein
FCGGRMLSNPPTKTKISIPTFVRNIIQSNLDEQKKTFTTGVGKVTENTVVRGDLTCT
DGLTTCQSDLTIQKINDIRPSTNSMVFDRGLNLSNPLSYKNSQGDITDYEMKSSGPK
QEVVTFQSGYKRTQIMMAQRAAHSSGYVGGSGVSPSGYVPWNKFQDTSQRTSGTEIYIDP
NDSTKLVEVNVKVPKLFRIISVIMAKHGSWLDNGTGADILLAANEYEGGGRINVTDLAM
TISRGSYYETRPLQVVCVITYAQNNGYFTFQNRAGGGLRVSSFWNIVALPYVE

CpN_0257 292136 292167

CT143 hypothetical protein
GVVMKRRNLKILPNASTPSTNVAENTGIKQNLFLDQATLVNNGVNDIENFLETRDLKV
ADTIISPCFETVGGGLSAESSQFATTLSKGLTSEDQGRVPKFTNVSDPQSPRDLT
YNYRNTGQCALNLYTYSSSQPTTVGKPIETVQCNPNPETYRISASAKIYDAVTRFFYI
QYKARGIYQVTIQRRESGHSGLDNPNLYNLNMGNNKTLTLCASDTRGYSGGHRTSIAV
TGTFTLTEIVATPHDYVWLFLETTIGLDIKSMSTCVIWFPPQANFAEVD

CpN_0258 292534 292133

CT142 hypothetical protein (frame-shift with 0259?)
CFSPCLRGSKFEKITLGGNTAIQLLAAGTYILTFTIGKRWGWNNGWGSIRLFEGKYTG
DGLGNTSTVYSGGYSTIGYLSAVYRDHSDIDDPNPNPDKYMNFLVFRNGDHSVIG
NYSFTLYFAGDKV

CpN_0259 293031 292441

CT142 hypothetical protein (frame-shift with 0259?)
IFYVFKRKTNYNFIEMTTNNODNNECYFKLDSTVDGDLASNQITFDKQAKGISSTETF
SVQGNATFKEKVSATGLTSASTYKLNATGPAPSSITIDKNNRSLNPLKPNPCDPVPA
VYRSPQYFFCAKPIESTFMFDGSSRYLPTIGDGSNTLYQSSKAGDVRFPVDQNSKKL
HGGTQYPNFLLOEPI

CpN_0260 294090 293548

secA-Protein Translocase Subunit
AYLDPSKRSCVEEDHVSKKINRNDLCPCGSKNKYKQCLKKEEQTARYTTEGKFKFAEV
LSASQGEADNCTKLQRLSQSLTSEQAAVKGQFHQITKNKVMSSKALKKAQAKEEL
VTEKLQOHNFYILNTGENLAPPMESTATLNQDNTNVCEDFIPTQEDPRISENSQKPEVE
D

CpN_0261 294272 295033

ydcO-P-Loop Superfamily ATPase
YGFNHPIIFPMFTLLNLPFWMKAGRIEGLVRKALYTHMTLANHKKIVVALSGKDSLT
LLMLKALGGRGFGLDIAVNIQYKYGAAYNKYLPTECDQICIDERTIPSPYAPETP
GYNCGAPARLLFOAKEIGACALFYHHRDLVQTALLLHLIKAFACMLPVIDMVIH
QVTLRLPLITPEFWRKPAKPIIPAVTRCTAVVIRKFAQDILKLEEVFLARINIA
LAFOHGGSSQKQI

CpN_0262 295055 295033

src-Src like Acid Phosphatase
LIPNINKEVYVLMNKRKILTLNDDITAKGMETVALLFANIGDIYIAAPQARQS

no robust homology present in Genbank/EMBL is of 11/7/98
KELFNLFEEFTANKETTASHELCTPHNGEGIGPPTLLCLLALSVLLLLGVVVALVASHVL

AAPGLLLVWCAASVCSMAIVSLMCLYKGGKPLIEPSNEEKIDPTKDLKIDPESLKPV
PVEGQSLPKERTVSPKAKIPSIVEDDFKPYVQISTFYHQNKVYKSPAEKMSLEKEIT
TLIVDFPRALLESSEKSGSLLRQISATFINKLPLPFLSRKVYKSLTACLPRGSLVEEYA
SSDLLILLTLTKPEPLAMVTOQLIAHLNLSKTERKRLTHMQKVLVSLNFWFYGMSLEEK
TEKIVADPNLLTDELKAHLEAGNIVQFLLSFQSSSEMRERFALFSDAQELPSAKDGSN
YVPAINGSEMYDFDKLVKSLSERLAFCEKIPSPSSWNTSSVASHYKDFSLLTFF
SNQSQSVTLQNPFLLELHENPKQCTFLKGLLEKAMPMSNWAALFRPMLMGLMCSGIARK
KELKI TAELHGVPFKEITQA IASGKILDLLQLHFLDF

CpN_0286 322149 322149 320650
mgE-Mg++ Transporter (CBS Domain)
SCRESKGMVGBQRNEEKLDTSSSGNLDLSRSHLDDLSFKLEKAFKLTSTDIHSH
DLSKIVIEYNPIDLAYAVSLPSESRAILYONLSCTAKVAFIINTDSASRAVFRRLSD
SEVCLAVDMPDEAVVWLDLIDRRYRILELIDSKALKIRDLQKHGRNTAGRLMNE
FFAFLEMTTVKDVASACIRSNPGIDLRLVFLDFKGLQGVVTDRLSIINPPEMSLKQIM
NQIEHKVLDPATREEVVDLVERYKIALPVDEENFLIGAEVVDVEAIEDIETIAR
MAGTTEVDVGYQCHVQVRLFLARPWLVLTLFAGLISASVMAYFQKISAPALLAIIFFIPL
INGMSGNVGVCSTILVRSMATGTLFGRRRRETIFKMSIGLITGVVLGLICGLVVMGL
FLGLNIFISGGGQLGVTVATGVGLASLTATTLGVLSPPFFAKLGVDPALASGPVITALND
IMSMIIFFLIAGGINFLFFN

CpN_0287 324230 322089
No robust homolog present in Genbank/EMBL as of 11/7/98
RRCMIRSPLEPFISSKRALNMLGLQDEFSCPEDVDVDFLSEIELLAQDDEPSGYLALSRS
LLMMTHNPKVVKRVI FYGVSYGLKHKMSIF IDPVLTYIDFLFEKLGISASDRSLCSAR
TCINFEYSQSGEMKFLSEVVDNFIQLKHKMLQKRLGWEHFRIGAKQEEVSLVAS
ASVYQAVGRSFIELVHKHLESLDLAGCMKCLALALDLSNNAHIHADYAKGLVVLGTROQ
KSLLIERGMEHFSKAI FLFSRSGQDTLAYQNYRYSALASVKLFDLTYYKHEHFGQAMNII
YQTVQAFPNLSGLVMWVGEILLIRSGVLNMSYKIEVLEKLASLQKKTNDPIALSGLLAT
GAILGLYLEEPLFKDSRHRLISAMRTFPGNSALVHALGVQVLCALYFNEDSHFASAI
SCPQSCLEWLDATGMQKFLDAYFSWG IKKKARLLKAVDVASRLCSLRPEAFLEWSD
RGLALCKLAETIDEAYKEIFLSESLHYQRAVLDLGRLEIELWQSSHYLLAEQLQSLF
HYDEATYLTIVKVDLTSSSRVKLILAAVLGLKGRLLQDTPAEEARELEIPLVEVYLEDE
NFLLLLGKYLFLFWKNKVCGLKARTYLEKATSLGCPAYYTLGKFYAVIKOVNKAWG
MVIRSAQYGVRITEAKWLNDDPYLANREIHAFREVVENQKRWLWGNKTEMKRN

CpN_0288 325785 324571
CT288 hypothetical protein
ISITIREPLFFGFECAKFNVMIMSCFNLTSTNESLRPISPKASFPKQGWQSYFRSALRK
HBSDTLSVSVCKVKNYDANFLVRLTIALAVVGVILFISIMLASIQGLTVITSWPLVTAA
ILEPITLITLGGMYILHRLGKLDVITSGVCIPIFFSRRCWVPISSHTLEKFEKHVSACS
LDJSTLSADGSGIAAVYQCPLFLRAFPFCGICAMPFFVALRLMYNLIRFLVVPYIIF
RMVEYHFFCKHLPEDDRIFYKDVAREMGRSLAALFKAPFYASACMIGAFYSLDPLAGRV
LMGSEVERDNDNVLARSVSLANEASLFRFEGGGGRKGLGQAFYLMCCQPSQVFLFD
KGEIVISGAHPSIQLPERRGLDTSGRYPHISVIPDSGNDSSAKNFIV

CpN_0289 325797 326996
CT289 hypothetical protein
NENRMLKKORSHYKKNLLLLLSILVGLGLSVQSPWIVYSAECIANTFLKFLRLLSIPL
VSCALGSTITSIONENTMTVLGRILYLTTLTVIAASIGLLFLFLLRPMITQDALATT
TKCNPLGLVDLSDTLPENIFKPFQCNVISAACLAVLGTASFLQEKHEKHFVNOFFNS
FESIFLNLARGGLKLLPIAMLGFSVILFKELKLDNLTMAEYLLCVIGANLQSGFVILP
ILGKINKVSPKVKAMSPALVTAFFSKSSAATLPLTMELAEDDLKINKNLSRFSFPLCS
VNMNGSCAAFIITVLVATVSNMGIISPLMSLGVFIATLAAIGNAGVPMGCVFLTSLSL
TSMNVLNLSGLILPFPYIDMIEITSLNVNSDCCVSVLS

CpN_0290 327027 328523
Na-dependent Transporter
RQALTMNKKHASFSSRLGFIISMIGIAGVAGNIWRFRVAAQNGGAFILWLCFLFLWS
ILIIILIELSGKLTAKKAPIGALIKTAGKFAWAGGFIITVTTICILAYSTIVGWSLSYFY
YAVSGKILHGNDFAKLWTSHYQSISPLWAHLTSLGLAYIVHGIKCNKILIPAF
FLCTIALLRAVTLPGAVQIGKQLFSCDKSCSNYKWIETALTONAWDTGAGWGLLVYA
GFKKKTGVVSNAGLTAICNNLVLSIMGIIFSTCASLDILGTQLQDGAASSIGITFI
YLDELFTRLPGGIYLTTLFSSIFFLAFMAALSMSIMLFLSQTALAEFGIKKPYISETLA
TIFAFVLGIPSAISLTFSSNQDTWGVGLAVNGLSIMYVGFPKLKEVAPAGDGL
RLNKAFDYIKYLLPIEGILLGWYFIEGLFPENGQWNNPISLYSLGLSVLQWSGLLIIL
WKNKQVLYLRSYNHEIL

CpN_0291 328658 329194
incB-Inclusion Membrane Protein B
EKHMSAPIPTPQELSDQITCLNVQVQVQSEELARENKGDIEGLKTLTAALTADAGIQPSAD
EYLSQTAAALILSAEKGPSGSGSTEGSVTVQSPCKFVALLAVLTLTIALIAVLA
CIIAACGGFPLLLSALNLYTIGACVSLPIIASTSVAILCCLTFVANSLIKPVITVTRTR

CpN_0292 329201 329836
incC-Inclusion Membrane Protein C
VKNTKNSDFMTPSIPFQSSGDASFIAEQPQPLPSTSEQSLVTLQTLTMMKHTQALSTELQ
QQRDLPTASIILOVGGAPTGGAGAPFGPGADDHHPPIPPVPAQIETETITIRSELQ
LMRSTLQOSTKGARTGVLVVTAIIMTISLAI IIIILAVLGTGVLPQVALLMQGETNLI
WAMVSGSICFIALIGTLGLIITNKNTPLPAS

CpN_0293 329940 332723
CT234 hypothetical protein
VMSQVRVLRLNPLNHGEEKRAFLLFLLGLVWIGCYGCTLSLAEGLFIEKLGSAELPKYI
LQCSLILCVLSSILYNLFKKHISATFIVLSLILCNFYLLISSIFAIDPPRSPFPF
YRIVTWSLTILSYTSFQGFVDQFNLDGKRKFCIFNAIIFLGDAGSGIIASLVHTTIGI
QGILILFTAALVLTFFIVFYVSKSLKSLSDHDLFDITGHPPPLSKALKLCF/DKYTFYL
LCFYFLMQLLAIAIEFNYLKIFETIQFASKEEFELVAHIOCKSLWISLGNMCFALFAYSRI
VKRLGVNNIILFAFLCLFSLFLFWTFKTLTSLIAVLAAMVYREGVYALDDNNQLLITGVPI
KNIRNQIRIVVESFIEPIGMLVWGLICFLSQQQYVFLCIIISLIATILVLCVGRYYAKAIL
NKLGAQALQTLRSMQDWIKSMTVQKQKQVELFLAHLKHPSERHQTFAPFHLNLNLSRSV
PLDLLAHMNLKLSLNLKLTIMVKSGLWAKDFLTLELLKRWTSIFPHFAIACAIIHLYFAE
HDLIIITHIAEDLYDTVGRLLAAIILTVRQEQAYGYRDLADKRLKELLNSDOPEDIVMG
UTILKLEKNFQNFPIILDFLNTKNEIDLIVTCALITVVRANIKFYCELLYFLRQCSHN
DEAQVYLLKTLTIALDIPVKKDLMTTQKLNTRKRYAAMIGLDEKVEAPAFVGLVLEDE
THNRCRIEAAKALCINDWLLKHAYIKVSKAKALFYGYKIHYYIQKRYPTYNLJLLA
NTIAGNYFAEVNFMLLJLILGSEMEHSGVLTALTSKNQKIKQAQALSELEKIDNHLPSL
LEFVYNQVIMCYSEKYYFKQGVILPLTELKLLNMMSNPOLNKLITAOQLKEELSYCDPDF
QGVNTIYNQHEHEDRFEESLETISPLSI

CpN_0294 333077 333502

CAMP-Dependent Protein Kinase Regulatory Subunit
IRNFMNLDRAFLKKKTIFQSLMDLLLTADKTTETIFKFGSNVFSIGQPGFSFYII
VEGYITISKEKLESPLNLKPLDCFGESLNNKPREYNASANTQVRMLVLGKQCLNIVE
ECPVALSFLLEYAKQIKFREP

CpN_0295 333866 333627
acpP-Acyl Carrier Protein
AMSLDEDDVIAIIVELQVDPKEVNESSFIEDLNADGLDTELIMTLEEKFAFEISEEDA
MLLDTYVNTYKFERCAT

CpN_0296 334712 334622
CT296 hypothetical protein
KPIRGMICMDITLVGKKVIITGSGRGIGLIGVKLFLENGADVEIWGLNEERGQAVIESL
TGLGGEVSFARVDVSNHGGVKKDCVQKFLDKHKKIDILVNNAGITRDNLLMRSEDDNQSV
ISTNLTSLYTTCSSVIRHMKARSGSIINVASIIGSAGQTNVAAKAGIIAFTKSLA
KEVAARNIRVNC LAPGFIETDMTSLVNDNLKAEWLKSIPLGRAGTPEVARVALFLASQL
SSYMTAQLTVVDDGLTY

CpN_0297 335724 334774
fabD-Malonyl Acyl Carrier Transcyclase
SRSNKDNFMKKRYAFLFPQGGSQVVGMDLYMEYFEVRELDFANERLGFSLTSMIFE
GPEDLLMETHSHQLAIYLHSMVVKVLQSRSSIQPSLVSGSLSGEYALTAVSDRISVLG
LELVKRRGQLNNEACNQSPGMAAALLGLPSEVIEENITSLGQGIWIANYNAPKQLVAGI
AEKVQDAIELFRDLGCKKAVRLKVGSAFHTPLMCVQADGLAPDIYALCMKDDSLPLVSHV
VGKSLVNTTEEMRECLARQMTSPTLWYQSCYHIESEVDEFLGPGKVLGALNRSIGISKP
ITSLTGTFQAEKFLSEV

CpN_0298 336742 335717
fabH-Oxoacyl Carrier Protein Synthase III
QYSEFYLWYFVSVNKKAAIATWAGSYLPEKVLNADLEKCMVDTSDSEWIVTRTGIKERRIA
GPETITSLMGAIAAEKAIANAGLSKDQIDCIIFSTAAPYIFPSSGALAQAHLGIEDVPT
FDCQAACITGVLYGLSVAKAYVESGTYNHLVLIADLKSFSVDYTDRTNCTVLFGDGAACV
IGESRPGSLEINRLSLGADGKLGELLSLPAGGSRCPASKETLQSKKHFIAMEGEKVFKA
VRMETAAKHSIALAGIQEEDLDWFVPHQANERI DALAKRFIEDESRVFSVHKYGMTA
ASSVIGALDELVTESIKLDDYLLLVAFGGGLSWGAVVLKQV

CpN_0299 336726 337415
recR-Recombination Protein
RKLIVVYSESLSNMLGPRPECKNKIHTMTRYPDYLSKLIFFLRKLPGIGFKTAEKLA
FFKLSWDSQGLKILGNAFHNVASERSHCLPCLFTLKESEKADCHCREERNQSLCIVASP
KDVFFLERSKVFGRYHVLGSLSPITGKHIEENRSLSKRSIETLCPKEITLAIADATLE
GDATLFLKQELQHFVSNIIRLALGLIGLSFDYVDSGLTARAFSGRHSY

CpN_0300 337768 340152
yaeT-Omp85 Analog
GRLLGMLIMRNKVLQISILALIQTPTLTSTFEKVKEGHVVDSTIITEGENASNKHPL
PKLTKRSGALFSQLOFDEDLRLAKEYDSVEPKVEFSBGKTNIALLHIAKPSIRNTHISG
NQVPEHKILKTLQIYRNDLFEREKFLKGLDLDLRTYLLKRGYFASVSDYSLHNNQKHGI
DVLKINENGGCKIKQLTFSGISRSEKSDIQEITQTKHSTTSWTFAGLYHPDIVEQD
SLAITNYLHNGYADAIVNSHYDLDDKGNLLIYMDIRGSRYTILQGVHIGFVLEPKRLI
EPKQVQGNPKIDYCPKIDWGAHKIKQYAKGYINTNVDLVFI PHATRPYIDVYVEVSEJ
SPOKSGILKITGNTHTKSDVILHETSLFPGDTNRLKLEDEORLRNTGYFQSVSVYTVR
SOLDPMGNADQYRDI FVEVEKETTGNLGLFLGFSLLDNLFGGIELSESNFDLPAGARNIFS
KGFRLRGOGGHEFLKANFGDKVTDYTLKWKPKHFLNTPWILGIELDKSINRALSXDYAV
QTYGGNVYTLINLEHLKYLGFYRGSQTSLEHKKRFLGLGPNIDSNKGFSVSAAGVNLAYDS
VDSPTPTTIGRGVTFEVSGLGQTYHFTKLNSNSIYKRLTRGILKIKGEAQYLPKFPYS
NTTAEVGPVSERFLLGGETTIVRGYKSFIIIGPKYSATEPQGGLSLLIIEEFQYPLIRQPN
ISAFVLDGSGFVLQYKISLKLDRSSAGFLRFDMNNVPMVLMGFQWPFRTETLNGEK
IDVSRQFFALGGMF

CpN_0301 340163 340762
(OmpH-Like Outer Membrane Protein)
IKDLSKEIFVVRKGFYWPFSIPKLQVIMKLLFSTFLLVLGSTSAAHANLYGVNLKRC
LEESDLQKKEETEELAMKQFVKNAEKIEEELTSYNNLQDEDMESLSDSASEELRKFP
EDLSGEYNQSQYYQIYAAALVYGFPKLKEVAPAGDGLRLNKAFDYIKYLLPIEGILLGWYFIEGLFPENGQWNNPISLYSLGLSVLQWSGLLIIL
WKNKQVLYLRSYNHEIL

CpN_0302 340766 341866
lpxD-UDP Glucosamine N-Acyltransferase
SKFPEFSMSAPVYTLKQLAELLQVEVQGNIEPTSGVEDISQAQPHHIAFLDNEKYSSF
LHNTKAGAIILSRSQAMQHAHLKKNFLITNESGPIFTOKCIELFI EPVTSGGFPGIHTPAV
IPKTIAREKNVTIEPVYVISOHAHIGSDTYIGAGSVIGAHSLGANCLIPKRVIRERVL
MGNRVVQDQAVLGSCEGFGYITNAGFHKKPLKHLGYIVGDDVEIGANTITDRGRFKNIV
IHETGIVQGVVAHVHIEGKHSIIVAQAGIAGSTKIEHVIIGQQTGTGHSIADHVI
MIAQTGVTKSITSPGIYGGAPARPYQETHRLIAKIRNLNPKTEERLSKLEKQVRDLSTPSL
AEIPSEI

CpN_0303 342982 341921
CT303 hypothetical protein
REQQGLHMDVSRKINRHTQFYVDSIDGVIKNFHKPSDEKSRDHEELEKLLTITKRIV
ASAQEQNRKTDKNMYLKKTOWLFPKNEELEQTKELFAMLTSMDDKIAQLFFYSFGCSS
DWEVFEITVCHLNSIGLOGLVLCCLGFEEQCEHVVTNKKLDLPLLGLTQVFNLSYYL
TYRNIOLLNCSMSSELGKELQDVLKQHVAFTLIFKEVIDLNLVFKLIQGLKRSQNIQ
ARIYDNDVPTLPSVSSSPIALRYSLANTIRGLALHVPSSLKFISSILNSTEHTAKALN
SGGECFIISNLDEFNLGMIMVQLLRTGKISPEILNKNIMKILMKRRVRSLYI

CpN_0304 343091 344158
pdhA/odpA-Pyruvate Dehydrogenase-Alpha
DQKPLPKRLPYKVMGDSAPYNIAQXTEKQTVERTILDYOPASCIKFLKQMVLTREFEA
RGEAYLEGLGVGYHISYAGQEAATAAIA NTYLDPPVWFSYPCHALAII LNIIPLOETAA
ILKAKETQYALARGGSMHMGPNFPIGPIGVYDILPAQAATIKYQQRKNRVSLCPIG
IGAVAZIVHTTLNFIYLIQLPLMLI IENNWMWMTGSLNRAVAKQPIAESQSSYDIRAV
TVNIFLBNLTLAGFAYRYMVDTESIVLX IIRFRGHTIGDNLVYKREMOCLPFR
DPLVIAKMLIRLEVLITFFEFQNIHQEYKTAVLAFANAKLQDDVTTIRGVYA

CpN_0305 344142 345117
pdhB/odpB-Pyruvate Dehydrogenase-Beta
KKDCMFKIKTLIKELARFALDEEMGPDNVI IIAFVVDYIYAAKVTYGLDKWIKPRV
LDAPISGAARFQIGGAALXILPILFPMWHTIIVALLQILGHAAMHMTKSKFVPI
VFRGPKAAAVQVQHSKVIKZIANIKELI IAIQYDAPYALIKALIRNNHVLLEN
ELIYNIKGEVITPEYLVPIKAIPIVQVRIITITTYGMYITPKFACILAKKPIWLSITCI

LDLRTIKFLDSTILSVKRTGRCIVIEEGHYFAGISSEIALITEHVFDSDAPPLRVC
QKREMPYJKILEQATLPMVNRILDTIEKVMR

CpN_0306 345136 346431
pdhC-Dihydrolipoamide Acetyltransferase
GKFVIGLKKMPKLSPTMEVGTIVKWHKKSNDQVSGDVIIEISTDKAILEHTANEDGWIR
EILRHEGEKIVTQTPAVLSTEANEPNLEELLPKTEPSNLEASPKGSEEVSPATTPOA
AGATFTAVTFKPEPLPSGLPVFKHVGTTNNLSPLARQLAKEKNIDVSIQSGSPGGRIVK
KLEAPAVLLLN

CpN_0307 348998 346515
glgP-Glycogen Phosphorylase
NGCIVEDFSSFDKNKVSVDMSKRAILDRLYLSVVGSPESASPRDIFTAVAKTVMWLAKG
WLKTNQNGYKNDVKRYVYLSMEFLGLGRSLKNSLGLDLVRKALKTLNYPDFHLVEME
SDAGLGNGLGLRLAACVLDMSATLAVPAYGYGIRVDYGFIDQIRIVNGYQEEAPDEWLRYG
NPWEICRGLEYLYPVRFYGRVIRHYTDSRQGVADLVDTQEVLMAYDIPIPGYGNDTVNSL
RLWQAGSPRGFEFSYFNHGNVYQAIEDIALENIISRVLYPNDISITGEQLRLKQEFVLS
ATIQDIRRYTTHICLDNLADKVVQNLNDHPALGIAEMMHILVDRELEPDKAWEMTT
VIFNYNTHTILPEALERWPLDLFSKLLPHRLIEIYEINSRWLEKVGSRYPKNDKRRSL
IVEEGYQKRIINMANLAVVGSAGVNSFSHSQLIKDTLKFEEYFPEKFINVNGTGTIR
RWIALCNPRLSKLLNETIGDRYIIDLSHLSLRSFAEDSGFRDHMKVGLKNKQDLTSRI
YNEVEIGVDPNSLFDCHIKRIHEYKRLNMLIRVYVNDLKENPNQDVPTTVIFSGKA
APGYVMAKLILKINSVADVNDQSRVNDKLVFLPNVYVMAEHIPGTDLSEQISTA
GMEASGTGMKFPALNGALTIGTMDGANIEAHEIGKENMFIPLGLEEQIVQLRREYCPQT
ICDKNFKTRIGLDELLEQGFNSNDKDLFPIVHRLHLEGDPFFVLADLESYIAAHENNVK
LFKEPDSWTKISTYNTAGMFFSSDRAIQDYARDIWHVPTKSCSGEN

CpN_0308 349213 349596
No robust homolog present in Genebank/EMBL as of 11/7/98
FTQENNMATVAQTPTQTPQPSVSHKATHRYCSWVFFKPIVLVSLGLLGLSLTTLGLVIA
SGVTLISLIGIVLAIQIVLAGIALVLAHNRQFKQARTAEINSMKMSAPAAATVQKQK
LEDYSSK

CpN_0309 350977 349595
CT309 hypothetical protein
FMRAWEEFLLQKEIGTNTVDKWLRSKLVCFDACLNYLEAQDSFQITWFEHIRHKVK
SGLVNNNNKPKIRVHTSVDKAAPPYKQKMQEKTAFTPMHGVSNPMTFSNFLVTPEN
DEBFVRLQEFKSPDENGGVTFNPILYFGPEGSGKTHMQSAISVLRSEGGKILYVSSDL
FTEHLVSAISGEMQKFRSFFYNIDALFIEDIEVFGSKSATQEEFFHTFNSLHSEGLIV
VSSSYAPVDLVAVEDRLISRFEGWVAIPHPVQEGRLSRFLMRQVERLSNRIQETALDFL
YALSSNNVKTLLHALNLAKRVMYKQLSHQLLYDEVDVTKLLKDLVLAAGSVRLTPLKIR
NVAYQYGVQSILGRSQREYVLPRQVAMVFCRQKLSLSYVRIGDVSFRDSTVISSIR
LESGKIEENSHDHMAIQDISKNLSLHKSLEFFPSEEMII

CpN_0310 353472 351049
60IM-60kDa Inner Membrane Protein
YFDELSLIFRVYQMNKRTLLFVSLIGIAFVGCCIFFGYNEFRSCKNLAEKQKRISEQTLA
AVESVGLSVASWTDVNGEHNKVAIVRVGDKLFLHNGEAAQSVSSGSEWSFVDHKKO
PDNHLALYRQCGSSFNPTNTGKVPFLPNHEGLPVLVVEFRNNKEPLVFLGEYAQGRISN
KDEHTFCTALVFWRSQSDYILGLYDSREKLVSLDLPITRAVIFGNDQDSAKSSDTANH
YVFNNDYMQIVSEESGSIENLFFASTNNKSIVNEIGFDRDLASEKSPALFPGLSK
LEDGQQAQNSIGGYPLLRGLLSDESKKLLPLEYHALNVSGRELATPVALRYRVLSTYP
HISQLESLDRSVQVVKLPENPEKPVFETAITLTKETEDVWVTSGVPEVEIMSNASAP
TIGSRVIRKKNGLSLDKVLPKVPKEPLARGVYQWILNSNGYFGIILTPLEISAGSGV
YISGSTAPRLTSAISPKNQLYPVSKYPGVETILLPLPKDAGTHRLVYAGPLAETPLKVL
DKTEPQKGENPEYLDISFRGVFAFATPAFAALLIIMKFFKLVGTSGWGISIILLTVFL
KLLLYPLNANISRMRRMQILSPYIQIQOQKYNEPKRAQWEIMGLYKTNKNVPTGCLP
LLIGLPLLIAMPDLKSSFLLRQASFIPIGWINLAPDVLFSWQTSWFIQNEFHLLPIL
LGIWFLQVQVTSLHKKGVPVTDQKQOQVMGNMMLIFTAMFYNPSGLNIYWLSSMILG
VVGQWINKILDSKHLKNEVVLNKKHR

CpN_0311 354453 353575
CT311 hypothetical protein
DMRAEMAVIWDORSKIVWSFEPWLSRLTWYGVFTVGFILACLARSYLALSYYGLKDHLS
FSKQSLRAVENFFYISILFVPGARLAVYIFYGVSFYLQHPFEEIIGIHGGLSSHGGLV
GFLWAAIFSWIYKKKISKLTFLPLDLCGSVFGIAAIFIRLGNFWNQIEVGTPTSLFWG
VVFSDPMQGVGVVHPVQLYEGISYLVVSGIYLYFLSYKRYLHLGKYVTSIACISVAFI
RPPESVYSHQSGKVLAECDLLTIGQLLSILPLFLGVALIITCLKARRHRSH

CpN_0312 354518 354976
CT101 hypothetical protein
CTMARNIKYFLILFPGILWISAGMKLLKATAIALDPLSSFFTYCLLSMVSWGLASLKHR
YLLSKTIRKQLSLSEFFSQKITWAIYIKQTFISRRFLIMVIMIAFSLVLRYSINPQAL
FVIRATVGYALIKTAIYAFISKLQNALMENPEN

CpN_0313 354957 355355
acpS-Acyl-carrier Protein Synthase
WKILKEISANMEIITHIGTDIEISRIEAIATHGNRLNRIPTAEQKYCLEKTDPIS
FAGRFACKEAVAKALGTOIGSVVAMKDIEVFKVSHGPEVLLPSHVYAKTGISKVILSISH
CKEYATATAIALA

CpN_0314 356285 355353
trxB-Thioredoxin Reductase
MIHSRLIIGSGPSGYTAAYASRALLHPLLFEGFQIGISGQQLMTTEVENFPGFPEGI
LGPKLMMNMKEQAVRFQTKTLAQDIIISVDFSVRPFILKSKEETEYSCDAIIATGASAKRL
EITGAINDEFWQKGVTAACVCGASPIFNKDLVYIGGDSAEALYLRGYVYVHH
RRDKLRAKAMARAOQNEKITFLWNSIEVKISGDCIVRVDIKNVQYQETITREAAAGVF
FALGHKPTNTDPLGQQLTDESQYIVTEKQTSKTSVPGVFAAGDVQKRYQAVTSAGSGC
IALDAERFLG

CpN_0315 356977 358716
ribB-Ribosomal Protein
MPKQAFYTWCKKILDNIECLTDVAEFKDLLEYTAHPITDIEEDNEIQCQALIKCTVV
DINKDFDVLVILKIEKVIEMSEFIDSGVILAEVEVYLDQANDEBKVVLGREGATR
ORQWLEYILAHKREIRIVKQITIRKVKQGLIVDIMEAPFLGQIDNNKIKNLDVYGVK
RFLIKINVERRNIVVSRRELEAPRTCKKAELEQIGIEYKRWVKNITDQVDFGLD
GIDGLIHITIMTWKRIRHPSEMVEINQELFVILVYKRGKVALALAKKHNWEDIEK

KYPFGKRVLVKLVKLLPYGAFIEIEEGIEGLIHISEMSVKNIVDPSEVNVKGEDEVAIV
LSIQKDEGKISLGKOTERNPNDTEEKYPICGLHVNAEIKNLTYGAFVELEPGIEGLIH
IDSMGKIKKVGHPSELFPKKNVSVAEVLGVDESKKITLGVKQLSGNPNWIEAMFPAQT
VTSVVTVKITAFAFVELQNGIEGLIHVSELSDKPFKIEDIISIGENWSAKVILKDPDH
KVVSILSVKEYLADNAYQDSRTLEDFKDSQSGPKERKKKGK

CpN_0316 358784 360121
nusA-N Utilization Protein A
GELRSKLVKVMKILVYVPMKEDQFORITICAFKALIAAKYKIPDANITV
DIPRTNHTVHKEKTVLQNDKFTDILVAREYDINQIDQYMDVFFVZDNFRITAH
AARQILIQKLRHAERDVILEYRHRVNETLGVVVKRFAGSNLIIDLGKVEAILPFRFY
KTEKKHIGDKIYALLVEQSENGGAELVLSRSHAEFVKQLFIQEVPELEGGSVIYKIA
REAGYRTKLAVRSDDPKTPDVGAFVGMGRSRVKNIRELNDEKIDIVNYSPTVELLQNL
LYPIEIQKALLEDKVIATVNDADYATVIGKRGINARLISHILDVELEVQRMSEYNKL
LEIQRLQLAEFDSPLDQPLEMEGISKLVQINLEHAGYDTIRRVLLASANDLASVFGISL
ELAYIKLBQVSKYSGESKVKDEKPEIED

CpN_0317 360045 362750
infB-Initiation Factor-2
SLLIRLSLKSANMEKVKLTNKLKIKNAQLTKAAGLDKLKQLAQGSSEAKSSSEKPS
AKEKSVKVALAATSTPTASAEQASPESTRIRAKNRSFSSSEESSAHTPVDTSEAP
VSIADPELELVVDEVDCEPSEVHPVAEVLPEQVLPETPPQKELEPKPVKPAEPKSVV
MKSFGPGTQKHNHLLAKTKFAPAKEEVVAGSSTKTPVASDKTKGPGTSEGGEQNNRE
KQFNANRSPASGPKRDAGKKNLTDFRDSKKSDESLKAPTRDRYGLNBEQEDRWKK
RSTVKKPKHYDEASIQRPTHIKISLPIITVKDLAEMKLKASEYIQKLFTHGMTVYVNDILD
SEYAFQITGLEFCTCIDIDYSEQDKLCLSNDRVDEIQSTDPKLVIRSPIVAFMGHVDH
GKTTILDSLRKSNVAATEAGAITQHMGAFCSTPVGDITILDTQGEHAFSAMRARGAECV
DITVLVVGAVDEGIEQKTELEIAEHAKAADIATVAINKCKDPNFSNETIYQLEINLLPE
AWGSGTIVMTSAKTGEGLSLELEMLALQAEVLELADKAPARAGLVISELHKGGLPVA
TVLQNGSLKGLGALVNDCYGKVKTMHNEHNEMLKAEAGPSIPVLTGLSDIPKAGDPFF
VYKNETARDIIEARSAGQQRFLAQKKRPNFDSMLQNKTKLMLIMKAVQGSIEALVSS
ISKISKSEKVDVEILTSVGEISESDIRLAASAKVLIGFHTGIESHAELIKSLGVRLV
FTVYHDAIDAIEIMTSLDPIAEKDEGSAEIKFIRSSQVGSITGYCIVTEGIMTRNKH
VRVLRNKEILWKGTLSSLKRVKEDVKEVRKGLCEGILLEGYQQAQIGDVLQCYEVIYHPQ
KL

CpN_0318 362704 363126
rbfA-Ribosome Binding Factor A
VMSYVMKLSIIHKYNLKYCYMTENRRIRKVNALLQEAIAKVLKDVKHPKISNLTWITV
RVSLKDLHSARVYVSMPHENTKEALEALKVSAGTAHRASNVVLYKVPPELHYLDD
IFSPQDYIENLLWQIQEKEKS

CpN_0319 363133 363879
trub-tRNA Pseudouridine Synthase
TIFFGNLTIKDMTMDLAVELKEGILLVDKPKQGRTSFSLIRALTKLIGVKKIGHAGTLD
FATGVNMLIGKRRFLRSLDILLFEDKEYEIAHLGTTDSYDCGKVGGRSKKIPLEEV
LSAAEYFQGEIQQLPPMFSAKKVQGGKLYEYARKGLSIRERHSTVQVHLQITKYEYPLH
FVVSCKGTYIRISIAHELGTMLGCGAYLEQLRRLSRGRFSIDECIDGNLLDHPDFDISPY
LRDAHNSL

CpN_0320 363824 364783
ribF-PAD Synthase
TPTISITFLPTYEMPMEIAYSLTSSFSVDSVTVGFFDGCHLGHNSLLTSYSGSSGVIT
FDSPHTVTLNHTKLINTKEERLQLLTQFPIDWLGLVTFDLNFANQSAEELFTLLHRLN
KCKRLILGYDSCIGKEQSSNTEALDTIGKPLGIEVIKIPYRMDGVNVSSEKAIROFLS
GNLECHRLFPGHPYASGKITEGSGIGGSLGATINLPREESLIPGVYACEIRYDSTTCQ
GVNMLGTPTFGRESLYAEAHIFSAENLYGKEVSIIPRKFLREKKFQSKETLIRAEIK
DILDADQWFAKGSFNYEGTA

CpN_0321 365900 364767
yohF-GTP Binding Protein
YSKHHYVIFIFRCLMSHTECGIVGLPNVNGKSLFNALTGAQVASCNYPFCTIDPNVGIVP
VDERLEALAKISNSQKIYADMKFVDIAGLVGASDAGLGNRLSHIRETHAIAHVVR
CFDDPDVTHVSGKNVPVEDIEVINLEIFSDFSAKNHSKLEKAKKREVGALLPLFD
TIAHLEKGLPLRLTLELPEQIVALKPYFPLTMKPMFYIANVDESSLPDMDYVAAVRE
VAAKENSRAVVICVRIEIEIVSLPIERLEFLMSLGLKSLGRLHVRDNDTGLISYFT
TGPQESRAWTVRGSSAWEAAGEIHTDIQKGFIRAEVITFEDMIECQGRAARELGKLIH
EGRDYIVQGDMLFLHNL

CpN_0322 366231 367328
yhcU-YopS Translocation Protein U
SNLNGSMGEKTEKATPKRLRDARKKQVAKSQDFSAVTFIVSMFTAFLSTFFFKHLGG
FLVSMKLSQAPTRHDPVITLFLYKNCMLMLITASLPLLGAVAVVGVIGFLVIGTFTFV
FKPDIIMLPNIENIKQFKIKTLIELIKSLIKIFGAALILYITLKSYSVLIETAGVSP
ITAQIFKEIFYKAVTSIGIFFLIVAILDLVYQORHNFAKELMEKFEVQKQFTEGNEP
KGRROIAQETAEYEDSSSQVHASTVVSNPDKIAVAGYMPKYPKAPWIIAMGINLRAK
ILDEAKYGIPIMRNVPLAQHLLDEGKELKFIPESTYEAIGEILYITSLNAQNPNKNQNT
NQPDHL

CpN_0323 367322 369460
lcrD-Low Calcium Response D
SFIMNKLNFVSRITLGGDTALNMINKSSDILALWMMGVLMIIIPLPPIVDMITINL
SISVFLMLVALYIPSAQLSVFPSLLITMTFRGLINISSRQILLKAYAGHVIQAFGDF
VVGNGVVGFIIFLIITIIQFIVVTKAERVAEVARFLDAMPQKMAIDADLRAGMID
ATQARDKRAQIKESLEYGAMDGAMKFIKGDVIAGVISLINVIGLITGVAMHGMDLAQ
AAHVYTLISIDGLVSIQIPSLIALTAGIVTRVSSDKNTNLGKEISTQLVKEPRALLA
GAATLGQVFGKGFPLWSFSLALIFVALGILLTKKSAAGKKGSGASTVGAAGDGA
VTDGNDPDDYSLTLPVILELQKDLKLIQHKTKSGQGFVDDMIKPMQALYQDIGIRYPI
HVRTDPSLBEQYDMILLNEVPYVRGKIPHHVLTNEVEDNLSRYNLPFTIYKNAAGLPS
AWGDEKALIEKAAIKYWTPLEVILHLGYFFHKSQSFLEQIEQVRSMEFMRSPDL
VKEVTRLIPQKQTEIFKRLVQEQISIKDLRTILESLGEWAQTEKDTVLTVEVRSKLL
YIUPFFSQGSAISVYLLDPEIEMIRGAKQVTAQSYLALDPEVNLILKSMRNTITPT
FAPKQPPPLTAIDVRRYVRKLIETEPEDIAVITVQRIILPEIRIQPLRIQIF

CpN_0324 369468 370688
CT124 hypothetical protein
VYVAHPHMAAGSGTQGLQGTQGLAALVFAAAKADAAEVAASQEGSEMNMIIQSQDLT
NHAAATRTKKKEEKQTLSESRKGEAGKALYVITTEEPDPTDLADKYA:KNGEISQQL
RLHDAI:IDGAGPDLALVQVETKEDALALYALDYLVVTPPQKLEALQIARNTHT
DUPRTAIGAKNILFAQYERADGLIV:DPDPIILHLYV:DTHTCDQLLSMLQDRTYQD
MAVYVGLMKMATELKRAQ:PYTPANGLVIMITTHHQAQVTS:SYDFESRVPILLDLKL

AEGLITQTDNLNFKVAETVYHKKIINDKPTASKVEREVRNLICDDVDVSVTVLNLFFSALR
QTSGRLFGCADKRRQQLGAMIANALDAVINNEDYPKASDFPKPYFWS

CpN_0325 370696 371148
CT325 hypothetical protein
KRIAMQNYEQLESAPLNTLAPDKNNCLIRFSDTHVPVQIEEDGNSGDLAVSTLL
GTLPENVFRERIFKAALGVNGQPSQSSIKGLGYEVTTQQLVLSILSMVNLNGEKLFEYL
KLFLSHAKIMESLRTNLPLDHLVGLIYVA

CpN_0326 371148 371148
malQ-Glucanotransferase
PSCFQNLRLRVNLYKTHKSPSAHAWKLICTSPKHGIYLPFLSIHTKNSCGIEFLDLIP
LISWCQKQGFVSIQQLPLNDTGEDTSPYNSISSVALNPLFLSLSSLPNDITPEVAKKLQ
DMHELCSPTSPSVSYTQVEKKWAFRETRYQCKCKSSLEGNNSFSELESERYWLYPYGTFR
AIKHHMHGEPINNWPKSLTDQENFPDLTKKHFDEVLFFSYLQFLCYQQLCEVKAYADQH
VLLKGDPLILISKDCDWWYFRDYFSSSSVSGAPDLYNSEGGQNMHLPIYNFSQLAKDDY
IWKERLRYAQNFYSVYRLDHIIGFFRLWIWSSSGRGRFIPDNPKDYIKQGTIELSTMLG
ASSMLPIGEDLGIIPQDVKTTLTHLIGCGTRIPRWRNWSDSAFIPDKDYNPLSVTLTS
THDSDTFAQWLNPSKEAKQAFKFLHLPQKTLTTETQIDILKLSHESASIPHNLFNDDY
LALCPDLVSKNQLRERINTPGTISKNNYSYRVRLSEELAIHKHFNQYIEKILTGL

CpN_0327 372927 373211
r128-L28 Ribosomal Protein
RIHRKNSRCKPLTGKRRRGGYSYTLRGIKKKKIGLKVTKTKRRFFPNMLTKRLWST
EENRFLKLIKISASALRHIDKLGLEKVLERAASKNF

CpN_0328 373220 374992
CT085 hypothetical protein
LKYREIFMSFLRRHISLFRSQKQLIDVFPVSPNLELAEIHRRVIEDQGPALLFHNVIGS
SPFVLNLTGTHKRVLDQLFQAPDNLIAVAHLISSTPKLSSLWKSRLDLKRISLGLKK
ARFRFPFVSSMSVNLDHLPLTSPWEDDGAFTLPLVYTESPTLTTPNLGMVYVRQFN
NTMGLHFQIQKGGGMHLYEAEQKQNLVSVFVLSGNPFLTSLAIAPLPENVSELLFATFL
QGAKKLYKKTNDHPHLLYDAEFLVVGESPAKRRRPGFPDGHFGYSLQHDFFPEFHCHK
IYHRKDAIYPATVVGKPYQEDFYIGNKQLEVLSPFLVPMVPGVRLKSYGESGFHALTAA
VVKERYWRESLTLALRILGEGQLSLTKFLMWTQDEVLDRFSVLETLERLQPDRLII
FSSTANDTLDYTPSLNKGSKGIFMGIGKARIDLPHGYQGGKIHGVODIAPFCRGCLVLE
TSLEDRICKSLHLHEDLKSPLIILADNLRETIOSEKNDFTTCAACDNLHSHF
ATHRPNNYFPFVIDALMKPSYPKEVEVDPSTKQKVSRWHAYFPNKETFYI

CpN_0329 375085 376146
Phospholipase D Superfamily (leader (33) peptide)
KMKRQKDKKIKICVISTLILVIGIPARAPRGDTFKTLKSEEAIIYSNQCNEMRKILCD
AIEHADDEEIIFLRIYNLSEPIQOQLTROQAQKNKVTIYQKFKIPQILKQASNVTLVEQP
PAGKLMKHQKALSIDKKDAWLGSAANYTNLSRLDNNILGMHSSSELCDLIITNTSGDFS
KQGTQKYFVLQDRKIAIAQAVLEKIQTAQTKTIQVAMPALTHSEIIQALHQAKQKRGHVDI
IIEBRSKSLTFKQLRLNINKDSTFTAPCTLHKHFAVIDNKTKLAGSINWSKQFSLN
DESLIILENLTKQONKRLMIWKDLAKHSEHTVDDKEEIKTEKSLPVEEQEAA

CpN_0330 376930 376202
CT093 hypothetical protein
FISIEMLLSRQLFSVLPSRFQDLHVYRFKESLKLQFMVTVGGIEIVVLAIEKEEDLRA
RKEFVRRKKNYLRIFRVLRSFDMRIIRFDYGLASQSIKADSRQNSPLVEKISEE
ATNEAIRALLAIGDRQEEKKRHYRLKGLQKQAKVLLSQLRHVLDLDFKLYCDSKKE
DQEKDEKNKQKRSIKVTKKKKGISLGAASQAIAAAAEAWIARNKGVLETAFTLYQKD
EEA

CpN_0331 378452 376701
CT082 hypothetical protein
IQRIIMAVSGGGGVQPSDDPGKWNPLQGEQAEQSPKLESIFSETQASSAAKQESLVR
SGSTGMATESQINKAKYKRAQDRSSTSPKSKLKTGFSKMRASVQFGMSGFRASRVSA
KRASSDSEGTSLPTENDVALKKGNRISEPMOGFFLDASGMGGSSDISQLSLEALKSSA
FSGARSLSSSSSESSVASFGSFQKAIEMPSSEKVNNAVVARLGEMVSSLLDPNVNTESS
LVREBAMATNGEMIDLSDLQEEVSTANTSPRAVEGKVKVSSSDSPKANPTGIPNSNTE
RAEKAEAKQESREQLSEDQMLAMRAMALGTGAAPQEVLSNSVSGSPSTVPPKFSGTL
PTGEGKDKSKHKSPGIEKSTNHTNPSPLREGTVKSAEVKSLLPHESMYRFPKDSIVSR
PEAVVKESTAFKPNPENSQNLPLIAVESVFPKESGTGALGSDAVSSSYHFLAQRGVSL
APLPRATDDYKELEAHKGPGGPPDLIYQYRNVAVEPILVRSQPFSSGSSRLSVQKGP
EASAVHDDGGGSGSGFGSDGRGSSGQKASRQKKGKSLTDI

CpN_0332 378676 378536
CHLTR T2 Protein
YLDNRIRVILPARQCTLLHLLAVLCPPISFFTQGVSPCVFFCFLDF

CpN_0333 379117 378800
ltuB
VDFFVVFVFMCKPKKSRTRDALAQEIQKSTEVLLKPKARIKAKNRRFLIAEKQTLKHR
AQEYDQVRLSLLDSQKKDQKVLIFNYENGFTVTDKDHFSKYSIRL

CpN_0334 379308 379823
CT079 similarity
TMSVHITPRKCFILCLISMTLPTLTPKALHILFSPYIVLCFYFCYKDKGLVLALGCCVL
CDLALGSGRGVFLLLYPLTALITHKAHLIFSKEAALVIVNMIFYGVFLLTIPMCALFK
HEVRWSIDVLMILPKCSFLDNLITSVIYLKCAINSEIKHMSIFRRRLVCY

CpN_0335 379808 380674
tdoD-Methylene Tetrahydrofolate Dehydrogenase
EIGMLLPQIPAAEKILQRLKEEISQSPTSPLAAVVLIGNDPASEVYVGMKVKKATEIGII
SKAHKLPODSTLSCVLKILERNQDPSIHGILVQLPLPKHLDSEVILQAISSPKDVGDLG
PVNMKLLKLNFGDGLLPCTPAIGIELLNYEIPLRGPHAAIVORSNIVGKPLAALMMQKH
PQNTCTVTLVHQSENLEPEIKTADIIIAALGALFLIKETMVAPHAVIVDQVTRVPADN
AKGYTLCDVDVFNNTVTRCAATITVPQGVGPMVAMLMGNTWRYQNF

CpN_0336 380569 381591
yoiL
KDKMKPCHHCHHRRRW.HHRYAHQVYAMMLPKFLVLLALGLAICQCKTTTIEGEQMTI
RFLIVLVTMLIAKEKALHQQIDRCFKHIDGIVNMNPNYSELINRAPADVPITLVEL
DFFLLGVYTYLKLDEGRFPVGPILKTLMLHLKLTPLPKDQVNDQHYKDMQWHLQFES
HFKTLIYNKHIVQIDLAANVKKYAVDNLGICNTGPNPNYVEMKIEIKTCHHPSRQWR
IPFAATFLDILDMALATPMMIKQKAVGKLYTHILDTKPKLELSYPIQSVSVVH
KQAYATAIDVIMLTPDKILAKQWADHHLIYTHIDGAS

CpN_0337 382141 381575
smpB- Small Protein B
IEEIFFPGNQGKRILILVLRPNKCFLLYWLFGPIMGEDLMAQKEIVSNRKLARNYEVITL
EAGIVLTGTEIKSLRDHOGNLGDAYIVSKGEGWLLNASIAPYRFGNIYNHEERRKRLKL
LHRYELRKGTEKIAQKGMTLIPGMFLSRGVYKVRGCCRGKAYDKRRITIEREKEREV
AAAMKRRHH

CpN_0338 382272 383375
CT339 hypothetical protein
FNMKTLIRNFIHMLFPLVAVONTPLIIVTHVLETHVHIAITATLQVTPRVTY
AKVYKGAISIPGKRFFQL/KELTEANLEISSAGEMAQITGSSGFRLLSMEKEDFPML
PDIOALRFLSPAEQLKTMLORTSFAVSGREESRYVLTGVLLAIANGVATIVGTGKRLAK
IDAEVTLKDSFSGEYIIPKAVEEIIKMSDEGEAAIFLDQDKIAVECDNTLLITKLSSG
EFPDPSMAVISTESNVKLDLHREELITLLKQVALFTNESSHVKFSFLPGELTTLTANCTKV
GEGKVSMAVYNSGELLEIAFNPFFFLDLKHSKDELVLGIDSISYNGIITDSASGLFVI
MPKRLHDD

CpN_0339 383405 384034
CT339 hypothetical protein
VITLPMFKICSLKLNFRNHSDEILSLAPKLNVAQKTNLLEALYVLSLGRSFRTOHLL
DTITFGSSHFLETFQEKDHLPOALSITYDKQKKICYNQLPIKTLSQLIGKVPILVSS
KORLLISGAPADRRFLNLLLSQCDNHYTLCLSYVHRALQQRNALLKSKQTSVASGMNS
WSNTAPTYSNGFSVVRNFQIYKPNFLGTT

CpN_0340 383842 384156
(frame-shift with 0339)
PLYPLLLVLSRSSSAEKCSLKKQANLNRGLWDEQLVKHGTYSIQRFLLCSQKLSDLKSEL
WSNNLKEQLALKFKSSLIKNSDISETAVAEFFHKQLSISLPRDLE

CpN_0341 384160 384495
(frame-shift with 0340)
SGSVGPHREDFTLTMQMPVSQFSSEQKHSLLAILRLAECLYLKQSHHVSPLVCLDDI
HAGLNRNVRVGLDPAITLQGLTLITSTHMHGELPKTSLVLSIENAVSEQII

CpN_0342 384619 385062
predicted OMP (leader (19) peptide)
HMKKTLITLIFLAVGNPLFSETSVIOTLPSGIGGLKETSQKQESVVCVHAFRLSYTSLKP
IARVLEKEHYDVFINWYETRKTFLKHAHLNRLKKIAELKPGVPINFTVHSIGGVIVR
VALAHPDCEAEKKEKFLSWLLRTQGLH

CpN_0343 384999 385595
(frame-shift with 0342?)
LPRRSQKRKAILMAPPNAGSTLARRYRCVFKVQFVGGKLGRLTYCTPKMLNVGKLPS
SLDVLILSGNRHSKFLPRLPYENDGKVCTIETKLDPHKAYVIHTSHTYITNRKSLYL
MKEFLKEGNTTPIIEHVPEAALEQTVMEKQKNSRLKPYNPQDIYVHCFSRPNLYGF
PKWWSLNQKNEINPEKLEK

CpN_0344 387432 385558
yaeL-Metalloprotease
SSRYMTIYFILAALALGILVLHGLHVVAKAVGMVESFSIGFGPALFKKRIGGIEY
RIGCIPFGGVYRIRGMERTKEKGEKIDSVYDIPOGFSKSPMKRILVLVAGPLANILL
AVLAFSILYMGGRSKNYSDCSKVGVWVHPVLAQECLLPGDEILTCNGKPYVDKMLTT
SLLEGLHNLLEIKRPGYLTVPKFEAIDVEFDPTKFGVPCSGASLYLTSNQVPLTKNSPME
NSELRNDRFVMDGTLLFMAQISQILNESYAFVYKVARNDKIFFSRQPRVLASVLTHTP
YLRNLDITQTPYSGSVLILRVQNHVRSITVQMSPOELEEVNSRDADKRFIASYHSEDLQI
LNLHGESHPEVAGPYRLDDPVQPRWIDVYSSESLDKQLEVAKKIKNKDKQRYVLERLD
AEKQKPSGLISLKLKLVRYNPSVVMVLSNITKESLITLKAIVTGHLSPOWLSGPGVIGVQV
LHTGWSVSEVLEFVWIGLISMNLAVALNLLPIPVLDGGYILLCLWEIVKRRRLNMKIVERI
LVPTFTLLIIFIFLTFQDLFRFFG

CpN_0345 388587 387436
CT345 hypothetical protein
LKVACLKHLAVLGTGSGIGRTLEIVRRYPSEFKIISMASYGNLRLFFQOLEEFAPLAA
AVNNEEVALSRKTIWILVILMSVFSDTIFLSSFLAVTLICMTALWGTILLISKQPLLS
ESLSHASYPGLLVGALMAQYVFLSQASIFWIVLPGCAASVFGYGIIVFLGKVKLHKDSA
LCFVLVVFVFAIGVILASVYKESPTLYNRINAYLYGQAATLGFLAATLAAIVFCASLPAL
WVWYRQI/VTTFDKDFAVTCGLKTVLYEALSLIFISLVISVSGVRSVGVILISAMFVAPSL
GARQLSDRLSTLILSAFFGIGSALGYSIVAFPTCAIIGQQAAPVTLPTGPLWVICAG
LALGCLLSPKSGWVIRFVRRKHFSKQDQEHLLKVFWHISHNLENISVRDQVCSYKY
QETPGPKPFRWRVQILEWGRVYKEDDYRLTKKRSSEALRLVRAHRLWESYLVNLSLD
VDEERFDSIRLAQVLEKQGGSSGFFNAANEVLRVRLCEEISWCDILKLTTLMECHK
VYACHSLEDILEVDGEARALAEI

CpN_0346 389690 388704
070-troD/ytdD-Integral Membrane Protein
KKGSLMALGSPSYGVVGFQFFSVFVSRFLSGSLFTGSLYIDDIQIIVFLAISCSGAFAG
TFLVLKRMAMYANAVSHTVLFGVLCVCLFTHQTLTLTSLGTLTAAAMATMLTGFLIYFIR
NTFKVSEESSTALVFSLLFSLSLVLLVFMTKNAHIGTELVLGNADSLTKEDIIPVTVIL
ANAVITIFAFRSLVCSFDSVFASSGLPIRLVVDYLIIQFSLACLVGAFKAVGLMALAF
LIPSLIAKVIKSIKSLMAWSLVFSIGTAFAPASSRAILSAYDGLSTSGISVVFVLT
MYIVVKFISYFRGYFSKNFEKISEKSSQY

CpN_0347 391078 389678
069-troC/ytcC-Integral Membrane Protein
TFGTNPEALSRKTIWILVILMSVFSDTIFLSSFLAVTLICMTALWGTILLISKQPLLS
ESLSHASYPGLLVGALMAQYVFLSQASIFWIVLPGCAASVFGYGIIVFLGKVKLHKDSA
LCFVLVVFVFAIGVILASVYKESPTLYNRINAYLYGQAATLGFLAATLAAIVFCASLPAL
WVWYRQI/VTTFDKDFAVTCGLKTVLYEALSLIFISLVISVSGVRSVGVILISAMFVAPSL
GARQLSDRLSTLILSAFFGIGSALGYSIVAFPTCAIIGQQAAPVTLPTGPLWVICAG
LALGCLLSPKSGWVIRFVRRKHFSKQDQEHLLKVFWHISHNLENISVRDQVCSYKY
QETPGPKPFRWRVQILEWGRVYKEDDYRLTKKRSSEALRLVRAHRLWESYLVNLSLD
VDEERFDSIRLAQVLEKQGGSSGFFNAANEVLRVRLCEEISWCDILKLTTLMECHK
VYACHSLEDILEVDGEARALAEI

CpN_0348 391815 391027
068-troB/ytgB-ABC transporter ATPase
FQWLLNVADETTFVSHNLCVNYEHAVALYHIFGLGVCGLTALIGPNCAGKSTLLKASLG
LIPYSGGTYFFNOKPKK/RQIAYMPQACVLDWFPMTVLDALMRYCYQKMMGRSS
DDPREAFHILERVGLSVADROIGQGLKXQQQPAFLAKALMQADLYLMDELFAIDMAS
PFGVGLVGLERDQGGKTIVVVRHDLCHVRQLFDHVVLLNRLICGPTDXTLNGDTIFQT
YQCEILLBQTLKLSRGKQGG

CpN_0349 391747 391749
067-troA/ytgA-Solute Protein Binding Family

31

NYGDAMEKLLVTDIDTITTHQSHLDKKVYERLYALHQAGWKLFELTGRITGVESGAAPLFSD
FDAPYLLCOONAGSAGWVGTGNNLLYSKSLPDLLEQLDQDMEGATALFVSGESAPGYDHY
YRFSPTEIQADLHEVYDPRYFENAKERELEFTETSLKDLAPFPSPAAAKVYFGLRDEVI RI
QKELERQEALETSVATMTLVMGSPDFRYALFLFTDKYSKGAALDRVNNILYDGKKPVMVA
SQDDANDLLETGDCGLVMGSPPEEMHYHAPLADPKNGILGAEWAGRYVYDNLMS

CPrn_0408 454090 454581
CT102 hypothetical protein
IPVLINQVLAALGKLVNLTATLSEETNTTMDPTMELMEEMKMSGLIMATIEG
GVWVAHPTDPLA CT102E CT102E CT102E
EPFPPDTIINHLDENIKQSSTASDTDFLSAPAKLRIRKKQTNDT

CPn_0409 454645 455127
CT260 hypothetical protein
MTTWTLNQNNLTFLKSSDEEPFLERESGLTYINIQANGNELPLFFVIRSEGEILQLICY
LPYQLVHSHKASTARLLHLNLRDIDIPGFMDEEQGLFYRLVLPLNGEIHDTLLRIYI
DTIKLVCDSPFSAIGLISNNMLDLERRLRQAQLQEQEQRNE

CPn_0410 455087 455933
dnaQ-DNA Pol III Epsilon Chain
DVLRFKSNKNKVMSSQTMDLVIFYDTETTTGTQIERDRIIEIAAYNSVTDSEFLTYVNP
EIPDEASKTHGTTDAWLSAPKPEPEAGFRKCGEEDSLVHNNDGDFPLLLGKECRRH
SLEPLNRTIDSLDKVAQKQPLDPPKHNQLVLRQVYGAENQAHRALDDVVILHKVFTSL
IGDLPPOQVLDLLQSSYHFKVFKMFPKGYKGQPLVDIPKSYFEWLENQGALDKPENKDI
KA
AIALHQQ

CPn_0411 455794 456609
CT262 hypothetical protein
RHQSRYSSTITSDMILTAAFSPCPNDIFLFRSFLKDPQFRPLLNQVTIADIETLNTLALQ
RLSLMKMSAALFPLVSDSYNDMLNDVGTNGVNSGPIVSLDPECSLDTLATPEGMTAHL
LCKLYLVIAKLPMPYDKILSAIDQKLVGGGALTHEERFSYDLQLTLRADFGEWRKRIT
FPLPLGCLAIKLYVPMATVDALTAALRKLSCISLCKOPITAGAKAVEYSKNGKVTVIHRFI
GTINKTEFOLSKTGGKALHLWKANECCEVT

CPn_0412 456515 457246
CT263 hypothetical protein
EPSTKKPFNNYLKLGKKLYICSGRPMNAVNTPKKILCIVADYREISPLIEQLDFTQINEH
LISYRCTQDHYLDLYIVHVGWSTAVNLALQSYCAQYTDYDLWINPGFVGACSPFPLGQCQ
TIEKIANLTDTDPVLSEDDPYIFDALPDSLPSKSLVTSPLYVHYGHFKHTFKLLDMGYA
IASQAAEHHPICPSFLKITSDYTPVGDCPFSLREIVSQKLQTQLVELLPELMEIRAPKLL
LPCP

CPrN_0413 459209 457227
msbA-Transport ATP Binding Protein
VPMKLLKAVLRHKNHVLVLGCSLALIGLITFSSQMEIFSLGMIAKTGPDAFLFLGRKES
GKLVKVSLSQGLDILENWAKISQKSEITLTVSDATTVAIEHGKSTASLTSLKSLFVRNYID
VSFRGLAFLDLCIVAFKAVTLEFQRFGLGVQVAIRVSRDLRDQYFKALQGLPMFTFHDD
IGNLSNRMTKDSALALVNSMLINQAPITFTLILTGKVLCSISWKFSILICVAFIPFIL
PIVVTRARKVIDAKRIHQSDSFSSVLYDFLAGVMVTKVFRTEKCFQFTKYCEHNNKISAL
EEKSAAYGLLPRPLHTIASLFFAPVVGIVYKFAIPPEELIVFCGLLYLYDPIKPGFD
ENSTIMRCQAAAERFYEVLNHLVLHLSQKEREIEFLGLSNTITFENVFSYQGEQDKHILKLN
SFTLHKGALGIVPGTSGSKTLLPDKSLPRLYEVSQGKILIDSLPTIENYKGSILNIACTV
LQNPFLFYDTVMNNLTGCKGDMEEAVALKRADEAFILKLPGKVHVSLEESGKNLSGG
QOORLAIRALLKNASILLDTEASALDAISENYIKNIIGELKGQCTQIIAHLKTLLEH
VDVRLYIENGQKIAEGTKLEELLQTCPEFLKMWELSGTKEYNRVFPDHKLVANPTDMAIT
T

CPN_0414 460203 459172
 acsA-AcsA Carboxylate/Transferrase Alpha
 LCLRLVICIKMLPIRGEIHMLPHEKQVEYEAKEAFEEKNKNLSLSSEIDQKLEK
 LRLDLKLEIKVLTPTWPKWQICRHPSRPRVTYVIEGMCEVFLGCDRTFRDDPAVVGGF
 VYIQGQRFVLIGQEKGCZDASRLRHNGMCLPBGFKRALRLKLAEKFGVLVFFVLDPDG
 AVPKGLTEERQGGWAKTNLFLSRLATPVYIEGEGCGGALGVGDVSMALSHSYV
 SVSEPGCAATLWKPDKNSEAAAMLKMHGLENLKQFGIDITVIEKPGGAHHDPAVLVYSN
 VREFIIEWRLRLDKLAPEELLEKRYEKFSRIGLYETITSESGPEA

CPn_0415 461522 460221
CT266 hypothetical protein
QDQTLRGLTLFVIFVYIWNCAFLIKLCVIMGLQSRLOHCIEVSNQSPFSGVQVKFYIAC
ISQKTFGLRSLKIFRYPHPLKLTHTDIARAVYLLMALEEGDGLGSFLNVOQYPSGAVELFS
GFGPWMLGYPAPAEHAFFGLLLQIAEFYESQAAYSKMSHQOALFDHQGSVFSLSLWGE
NSRLLLEKTTLSQSFLFOLQMGHPHEYSLELDPALGFWMORTFSQAFVAASGCOSSQAGE
SSGDVGYIAPPCSGDISDQYPCGCCIGAPFVCKQSHOTTEISFLTSTGKPHPRNTGS
KVSQGVYHLPIRCKITISDKQYVHAALAEATSMFTFSFCCKNGCQVQDVPLRSGSLSD
VYKSGPGNDIMILGENDAINISVSPYMEIFALQGKEKFWNIADFLINIPYKEEGVMLIFEK
KVTSEKSGREFTKMN

CPn_0416 461871 461557
nimD/ihfA-Integration Host Factor Alpha
EALSNMATMTKKKLKLISTISQDHKKIHPHFVTVIQNPLDKMTDALVKGDRLEFRDFGVQLVQ
VERPKPVGGRNPKNAAPVPIHPARRAVVFTFGKRMKRLIETPNKHS

CPN_0417 463047 462244
 CamiA-N-Acetylmuramoyl Alanine Amidase
 KRGKMKLTKYLTNKLQRLSMISRLFRVYSLPMSKQLSFFALCVLGSHPIFAQTNPNPQVRV
 KSEVIFIDPHGGGQDQSTALKHELHYEKSLSLTSLATVQYQYKLRMGYKPOLTRSSDVVD
 LKRAVLVLRGQKQDVFISIHNCNHSAAAGTFVFNKYG/SPTRNRMSEVLGNKILNLA
 MEKNGLKSRGLKLTANFVVIRDTSPMAVLVETQFLSNSPERAALQDARYRMHVHGAIEG
 YHNLFGQSPFOKPKNTAKIRKPOIAEN

[illegible]

IP:192.177.81.100 V:1.0.0.0 R:0.0.0.0 P:0.0.0.0 LMAQ:VVERYGFA77T SDNF:CEPTEPEDIJNEICDG
FY:KNZFIEIDKQAAITTALE:IASDFDITVLIAAGFCHEATYQIFKHQITVAFDDPQTVCEVLA
CV

CpN_0419 466977 464876
 bbp3- transglycolase/transpeptidase
 QLFPTNTNINIPQKVKSVFYPMYSYRKRSTLVLGVFALYALLVRYYYKIQICEGDHAAE
 ALGQHEFCVRDPRRGTFFANTTVRKQDGLQPPFAVDITKFLCADPLAIPCHRDDEII
 QGILQFIEGGTYDGLSKLDKKSRYCKLYPLLDVSVHDLRLSLWAGYATKHLRPTNALFF
 ITDVQSYFPFKLLQGVHLTLREIKDEKTKAFPTGGMEAYFNHILEGDVGERKLLRSL
 NRLDTNRVTLKPKDGGDIYLTINPVTQTAEELERGVLEAKAQGGRLILMNSQTGEILA
 LAQVHFDHPTHTDGLNKKKFDHFFVTVVATERGIMKFTTANALCANFEASQYGC
 PFIDHDEEDHATPTHTNPPHILNKKKFDHFFVTVVATERGIMKFTTANALCANFEASQYGC
 VAWYQKLLALQFGKRTGIELPSEASGLVPSPHRFHINGSLEWSLSTPYSLAMGYNILAT
 GLOMVAAYILANGGYAVRPTLVKKIVSAGEEYHLPTKEKTRLFSEEITREVRAMRFT
 TLPQSGFRASPKHSSAGKTCITTEKIMHCKYDKRRHIASFIGFTPVESSEGNFPPLVML
 VSIIDDEYGLRADGTYNMGGRCAAPYLPASVADRTLLVGLILPDKKLRNCDDEAAALKRL
 YEWNRSRPGQGTGTR

CpN_0420 467120 466824
 CT271 hypothetical protein
 KSPFMNKSRLRCLCCFCGSLFFYFYINKNSLTKLRLEIPCLSVRLRQLEQQNISLRF
 LIDKIERPHLMIEAALPEYQYLEEYPSSEISLSEYLP

CpN_0421 468007 467108
 yabC-PBP2B Family methyltransferase
 EILMSERAIHIVLVEECLALFAORPPQTFRDVTLAGGHAYAFLEAYPSLTCYDGSRDRL
 QALAIKAEKRLTFFQDRVSFASHAFEDLANQTPRLYDGVLDLGVSMQDLTLRSGFSFQ
 GEKEELDMRMQDTQELCASDVLSNKEELGRIFREYGEQPKWSAAKAVHFRKHKKIL
 SIQDVKEALLGVFPHYRPHRKIHPLTLIFQALRVYVNGEDRQLKSLTSAISWALQOGR
 VIISFCSSSEDRPVKWFKEAEASGLGKVIITKKVIQPTYQEVRRNPRRSRAKLCFEKASQ

CpN_0422 468233 468784
 CT273 hypothetical protein
 GLAMVEIFNYSTSIYEQHASNNRIYDFRKEIQMEGISIRDVAHQAILDMNPKPSALTS
 LLQTNQKSHWACFSPPNNFYKQFSTPYLPSLSPDQDDEIEKISSFLVLTRGKFYQ
 RSQITPFLSYKDEEEDEDEPDEDDDDPRVQQGVLLKALDGVKSTNVMIDYVISRIFQ
 FVQG

CpN_0423 468788 469216
 CT274 hypothetical protein
 CMLDNEKAKILGWGDELEELISGYSFLRQGHYSKAILFFALVILDPLSIYDHTQLGG
 LYLQIGENSQALAVLDQALRMGGDHLPTLNLKTKALFCLGRTEATATYLSSCPAPAI
 ANDAEALLMSYKATKKNAAVLR

CpN_0424 469528 470961
 dnaX-Replication Initiation Factor
 SRQNEIFSPSLMGWDCIWESFINKESGMLTCNECTTWEQFLNYYVKTRCSKTAFENWISP
 IQVLEETQEKIRLEVNPINQVNYLLDNVYKRDLCFVPLDVHGEPALEFVVAEHKKPSAFV
 ASDKESNEGISEVEETKDFELKHLNLYRDNFIEGSPNQFVSAAGVAGKPGRSYNPL
 FTHGGVGLGKTHLLHAGVHYVREHKNLRHCITTEAFINDLVYHLKSKSVDKMKNFYRS
 LDLLVDDIQFLQNRQNFEEFCNTFFTLINLQKQIVITSDKPPSOLKLSERIARMEMG
 LVAHVGPIDLETRVAILQHKAEQKGLLINEMAFYIADHYGNVRQLEGAINKLTAVCRL
 FGKSLTETTVRETLEKLFSPKQKISVETILKSVATVFOVKLINDLKGNRSKDLVLAARQ
 IAKYLAKTLLTDSLVAIGAAGFKTHSTVLYACKTIEHKLQNDETLKRQVNLCKNHIVG

CpN_0425 470965 471564
 CT276 hypothetical proteins
 FRGCPMFRRTGKPPFEDVQTYLEETSSSPSSPYSRSERPETPPSLFDNPKASEARPLN
 HNTLEESSLPQWSTPTRESLLPLEEPITTLGEGVTFKELAFERLLRIDGTFFEGILVSK
 GKIEGPKGVVKADIQLOEAIIEGVVEGNITVSGKVELRGGAIIKGDQIANTLCVDEGVR
 ILGLVLAITAGTIDHSERERDL

CpN_0426 472111 471536
 CT277 similarity
 MVLFSLFLPKLCYGCQAPGAYFCSNCKLEKLLVEDREGRCLHCFRYLGSSETRLCSCQSPS
 SQGAFSLVLPQSTALSUYARACGKRPALQFFSKSIAFELASLDETSPCIAIYITSTISR
 KIVVEKALEKLRILPLWPLPKKRQIEKLPKGEICFLSAYPLSQKWMQTIVGGSASPL
 VSISEFLSQNDQ

CpN_0427 472153 473715
 nqr2-NADH (Ubiquinone) Dehydrogenase
 AVCYVFERVEASTFLSITMLKKFINSWLKLCQDQYQRFTPIVDAIDTCYEPITETPSKP
 PFIRDSVDVKRMMLLVIALFPATFVAIWNGLQSIYVSSGNPVLMEQFLHISGFGSYLS
 FVYKEIHIVPILWEGKIFIPLLTISYVGGTCEVLFAVVRGHKIAEGLLVGTGLIYPLTL
 PPTIPYMAALGIAFGIVVSKELFGGTGMNINLNPALSGRAFLFFTFAKMSGDVWVGSNP
 GVIKLSLKMNSSTGKVLIDGFSQSTCLQTLNSTPPSVKRLHVDIAANMLHIPHVTQD
 VIHSQPSLWETHPGWLDNLTLTQLQFTVTAPEAGGLGLLPTQFDSAYAITDVIYVIG
 KFSAGNLPWGNIIIGSLGETSTFACLLGAIPLIVTGIAWRMTAAFGIAFUTGWLKPFIS
 VLVGQNGAWAPARFFIIPAYRQLFLGGLAFLGVFMATDPVSSPTMKLGKWIYGFIFGFM
 IVIRLINPAYPEGVMLAILGNVFAFLIDYFAVRKYRKRGV

CpN_0428 473719 474681
 "nqr3-NADH (Ubiquinone) Oxidoreductase, Gamma"
 NMSKSGSKHTVRINQTYIVSFILGLSLFAGVLLSTIYVLSPIQEQAAATFDRNKQMLLA
 AHILDFKGRFOIQEKKEWVPATFDKKTQLLEVATKVKSEVSYPELEYAERFVRPLLTDA
 QGKVSFEEKNLNPIEFPEKYQSPCCQSPPLFYVILENTSRNENSGADVAKDLSTVQ
 ALIFPISGFLWGPINHGVLGVKNDGDTVLGTAWYQOQGETPGLGANITNPEWQEYFGKKI
 FLQDSGTTNFATTDLGLEVYKGSVRTTLGDSPKALSAIDGISGATLTNCNVTEAYVQSL
 ACYRQLLINFNLTHEKTKCE

CpN_0429 474666 475319
 nqr4-NADH (Ubiquinone) Reductase 4
 KENPMATCKKYSKYFFDPLWNNQILIAIGLICLALAVTTTQVATITMGIAVQIVTQCS
 HFPVGLLRKFTPDVRMTQILIIILFVILIDQLKAFPPDINKTLVYFVGLIITNCLVM
 HKDEGLARHVTIPAFLDGFAAGLYKWWLVIGVILPFGOTLMQFRIFOPVYASST
 HPGFYNLHMLVAPSAFLLAHIMLVNIRDKYPPR

CpN_0430 475323 476013
 nqr5-NADH (Ubiquinone) Reductase 5
 FMNIGAYTWLWVGLLQAAFIQNIILLANFLGMCYIACITRVVTANLGMVALVLTVT
 CTHWPFHAFITPKALTWSPSLASVNLGFLCIIIFIVYIAFPQILLELLKVRNLY
 LSGFLPLILAVNGLAVLWGLTPRVSYPFMHITDAGQWMLAVLATIKREKLY
 SDIPKHAQMGITPTNGLTAMAFMLTIDIDIKAKIKRAHLCDEVTENTNPKERS

SKHQPSISKAPTQRRSL

CpN_0431 476489 476151
 No robust homolog present in Genebank/EMBL as of 11/7/98
 KIMTLPKYVPRSRQNPDTLTLKRYSSVLLHSSENSLSYRIFAKVLAILLTSLAVAFVT
 LFSCEGSQLRLCALYIGIALACVLLTIVVYICIAJKIATACKKKPPSISRIEIV

CpN_0432 476837 476514
 No robust homolog present in Genebank/EMBL as of 11/7/98
 GLEFLTGSAPLQJLWIWIAASITLGLMLVCAWRYALNADLETKVAHES

CpN_0433 477327 476929
 gcsH-Glycine Cleavage System H Protein
 RTFRILYGTLYRTGSRKVMWYSDYHVVILPVHERVVRGLTEKMOKNLGAILHVDLPSVG
 SLCKEGLVILESSKSAIEVLSVSGEVIDINLDLVNPPKINEAPEGEGWLAIVRLDQ
 DWDPNLSLMDDEE

CpN_0434 479471 477276
 CT283 hypothetical protein
 RFWRIYQQDLFCRLCRDPAWFSLLSFTLRFYCLGRGWTLSSFFYKHQKXFIGIVIAV
 VCSGIGVGGWGRFSRKGSAESTSRRTVFTTASGRYVEKDFMAMKKFFAHEAYPTGNPRA
 WNFINEGLLTDYFLTRVGEKFLKVVYHPGEKIFSKKAYQPYRRFADPFISSEEVKSS
 APDOLLKLVQFQIENPISKGEFLARAKFLERREFHYVLRQMLEYRQMFALPPDEAL
 SRGKDLRLFGYQTIQDWFGDAYLSAAVELLIRFIDEQKQVLRPSKQEARDDFYDKAKHA
 YTKISKONKESLGFEEFVNSYFQFLEISESEFNNMYRDLILCKRALLLQGGVSFDFQPL
 TTFVYQGGKQIYVEFFRLPKKEYSFKTKQELKAFVYKLVSLPKSDSLQKNEILPITAT
 KAKEPRLVGRRFIDYKRVALQDLAATVPMVVLWMOQNSSEHFOEILQOFPDVECTQSYK
 DFQHLKALRDLKLSLFRKEILRARPERILQSLQVQPKQSQEVLLSAGKMLLPISDQ
 QLAQLLENVLDLYSQAETYYTIIIVNSFEKEEVLPYREVLKRDLASQLLTHSHGLVD
 MERLESALRTTRYPEEGEASLWQRLWKVVENHRLGRHLEGFSWSLDRSLKTFIRSGDKEL
 PQEFDRIFSMKVGDYSSVFMSPNEGPCYYQCLSHLLYDRPASVDKFLAKSQLDEELLGS
 YMERFIEQGVVR

CpN_0435 480908 479475
 Phospholipase D superfamily [uncleavable leader peptide]
 GVMSRLRLAALGIFFILLVNSVSASKTIVASDKKGVVLYVNSVEFAQQILDIDH
 ANFVLELCPMTGGRTLKEMVDHLEARMDLVPELCSYIIIOPTFDAEDQKLKALKERH
 PNRFVYVFTGCPSTSLANVIMEMIKLSIIDGKYCILGOTNFEFMCPTGDEVPEKVD
 NPLRPLVSGVTRPLAFRDQDILMRSTAFGLQREYKQFAMWYIAHMMFIDNPEFAG
 ACPLPLLEQAETVPFGFDKHEDLVLDSSKIRIVLGGPHDKQPNVPTQYELKIQGARS
 SVKLAHMYIIPKDELLNALVDVSHNHVHLSLITNGCHELSPAITGPAWGNRYINFALL
 YGKRYPLWKKWCEKLPYERVSYIEFAIWTQLHKKCMIIIDDEIFVIGSYNFGKSDAF
 DYKISVIESPEVAANKVFNKDIGLSIPVSHGDISFWSYFHSVHHTLGLHLTYMPA

CpN_0436 481633 480902
 lplA-Lipotease Protein Ligase-Like Protein
 FYCYMKVRIVDSGKSSAASHMAKORDLLESQDGLLILHLYEWENPCSLTYGHFMRPEK
 FLSNVADLGLDAVVRPTGGGFVFKQDYAFSLMSATHPSYSSSVLENVHTVNSFVAKV
 LEKVFRIQMLAPEDENSSSRDSGNFCMAKTSKYDVLFGDKKIGGAQRKVQVQGLFQGS
 LFLSGSSSEFFQRLKPEVLEETIEQIQIHAFPLVGLLEADEVLEQARQVKAFFIKLFC
 GEG

CpN_0437 481810 484350
 clpC-Clp Protease
 VFMEKFTNRKQVILKAKKEAQRNLNHNLYGTEHILLGLLKGQGVAVNVLRNLGIDFTD
 ARQEVERLIGYGEIIOVYGDPAITGRVKSFSFANEASLEHNYVGTHELLGLIHQSD
 SVALQVLENLHIDPREVRKEIRELETNQLQSPSSSSSSSSSSSSSSSSSSSSSSSSSS
 DKNEKLSALKAYGPDLTVMVRESKLDPIVGRSSVERLILICRRRNKNPVLIGAGVKG
 TAIVGEAQKIIILNEVPDALKRKLITLDDALMAGTKYRGQFEERIKAVMDREKHCNI
 LLFIDELHTIVGAGAEGAIDASNILKALARGEIQCIGATTIDEYRKHIEKDAALERR
 QKIVVHPSPVDETIELRLGLKKYEEHNVFITEALKAATLSDQYVHGRFLPDKAIDL
 LDEAGVRVNTMQPTDLMKLEAETENTKLAKEQAGTQYQEKAAAGLRDEEKLRLERLQ
 SMKQEWENHEEHOVPEDEEAVQVSLQTGIPSAIRLTAESKLLKLEDTLRKRVIGQN
 DAVTSCIRAIRRRTGKIDPNRFTGSLFLGLPTGVGKSLAQQIAIEMFGGEDALIQVDM
 SEYMEKPAATRMGSSPPGYVHGEHGLTEQVRRVGVVLFDEIEKAHPIDMDLMQLIL
 EQGRITDSFGKRVDFRAHIIIMTSLNGLADLRKSGEIGFGLKSHMDYKVIQEKIEHAMK
 HLKPEFINRLEDSVIFRPLEKESLSEITHLEINKLDRSLKNYQMALNPDSVISFLVTKG
 HSPEMGARPLRRVIEQYLEDPLAELLLKESCRQEARLRLATLVENRVAFEREEREEAAL
 PPHLES

CpN_0438 485455 484334
 ycbF-PF-loop superfamily ATPase
 NLTLPMPQVREIMQTVIVAMSGVDSSVAYLPKFTNYKIVGLFMKNWEEDSEGGCLC
 SSTKDYEDVERVCLQDIPYTVSFAKEYRERFAELKEYSLGVTNPDILCNREIKFD
 LLQAKVTELOGDYLATGHYCRNLTELOQTQLRCCDPOKQDSYFLSGTPKSAHNVLPFL
 GEMNKTEYRAIAAQAALPTAEKDKSTGICIGKRPFEKLEKFLPNKTNVDMWDTKEIV
 QHQQAHTYITQORGLDLOGSEKPCYVVGKNTENSIIYVRGEDHPOLYRELTAELN
 WFTPPKQCHCSAKVRYSPDEACTIDYSSGDEVKVRFSQPVKAVTPQGTIAEYQDGTCL
 GSGVITVPMIPSEG

CpN_0439 485523 486077
 No robust homolog present in Genebank/EMBL as of 11/7/98
 IISSNLP/LFVSSSTLNGVFPSSLPEESADLFTNKEIVALGEGKNVFLTHSIIMHIAAIT
 IIVLVALAGTAICLCQYSQSLIIAIVGIVLTILTLCLLQALVGFIFIRQLPOQLHLLT
 QFIREKIPRESSLQVLTNAQRKTTQDTLKLIEELCDLSQKEFKLQSTLYQKRFELSHKNE
 KTNQN

CpN_0440 486081 486740
 No robust homolog present in Genebank/EMBL as of 11/7/98
 LATIPGRMATWAPSVPRTPLCHATEVNLHFNAYITOPHIPAFAFWETFRSLKTH
 LITATLALLTLAGTIAGYAGYTYNMIIGYGLGIVITGLIALLAIIPLKKNQGTGKL
 IDELQVIGISGVORYGLMEFTIKVHLDELITQNGQETPLINEIAEKKEIQNLEL
 PITEXCHLAQKQPKRKSQKSPMPKIKHKNPVIITDQ

CpN_0441 486876 486838
 CT007: hypothetical protein
 RWRILGAKRLIHAFAAGHVLTPITVQDAGYIDRACFNPPDPPEAQVYLVKND
 AKPKKLPHOTIGYRQYDITFLCTLPITPHEGLSTGYIGADLWKLGLLITDPNGLG
 WATQDQTPVNVVLLSAYTLGLKNWOWHILEKLVNPKHILEYGLYQVLSQKYQAT
 EKIAIPQVINTGLHQRKAWLWVYKATATLTLPIYVNPISIDYRTPVGNALAY

RLTRFRKKLHKHNLISPGIFEEYAREICANVKLTPTWGLSFIKGFYAWIGNDISIADDDH
NNNKTSHTEKTSIAFFEGGSAMNMF

CPn_0442 487942 489528
CT006 hypothetical protein
NILPQTMDFGNKICKQGGQLYLNGIFPERILARKLKNCAKSYPRALTITIEVLVSSVLGAL
KVTLPICASTYAALTLPRLALFNALIKTKSCQGLHASYAMAWLLHITIAVIIGLVFSLVFI
PPPVFVLSLGLMLGVTTVTLPFG/HKNLPFPPEPPSRPTPPPPFAADYVPLISESYFD

[illegible]

Cn_0444 490266 494507
 pmp_6-Polymorphic Outer Membrane Protein
 KAKFPTMYKYSLLPWLLTSSALVFLSHLPMTAANTDLSSSDNYENGSSGSAAFATKETSAS
 GTGTHMLTSDVITNVAITPADKSCFTPTGALTSFGVADGSLVLQTIALTHDGAANNST
 TALSFSFGFSSLLIDSAPATGTSGGKAICAIVTNECGTATFTDNASVTLQKNTSEKDGAAV
 KYSISDLAKTTHAALLDQNTKGGKALSTANTVQGNSTATFTSSNTATDKGQNIYSK
 EAKSTDLANTGVTVFSKNTAKTGGAWSSDNLALTGNTQVLFEQNKNTGSGAAQANNPGC
 GGAACCCYLATATDKTGLA1SQNGEQVPSFTNTTANGAIVATKCTLDGNTLTFTDQNTAT
 AGCGCAIYIETEDFSLKSGTGTFTSTNTAKGALYKGNSSLTGNTNLLFSGNKATGP
 NGSNSANQCGCGGAILV1IDSGSVSDKGLS1ANNOVEFNTSNAATVSGGAIYATKCTLG
 TGTSTFDGNTAGTSGAIFYIETEDFTLTGSTGTFTSTNTAKGALYKGNSSLTGNTN
 LLDGSGNKATGCPSPNSANQCGCGGAILV1SFLEASAVTSGKGLVTEDENVLSGNTATVSGG
 IAYATKALHGNTLTFTDGNATETAGGAIFYIETEDFTLTGSTGTFTSTNTAKTAGALHT
 KGNSTSTKFNKALVFGSNSATATFTTDDQCGCGGAILCNISSDIATKSLTLTSENLSF
 LNTAKRSGGGIYAPKCV1SGSSINFGDNTAGSAILYKNSLTANGPVSTNTNSNGG
 KGGAIY1ADSGELSLADGDITFSGNRATPESTPSNLHLGAGAKITKLAAPGHTIYF
 VDTIMPTEAPASGGTIEELVNPVIVV1VPPQPKNGI1ASVPPVPEAPANNQTI1VFS
 GKLPSQDAS1PANTTITLNLKINLAGMNVVLKEGATLQVYSFTQPQDSTFMDAGTTLT
 TLTNTDGS1DLKLNVLNLDLKGKMT1IAVNSTSGGLK1SGDLKFHNHNGSFYDNPGL
 KANLNLFFDLDSSTSGTVNLLDDFN1PSSMAAPDYQSGSWTLVPKVGAGGKVLVIAEW
 LQVTLKPEKRLATVPLNSLNAVFNH1SQEIALTGKSDAPSHGV1GGIGNAPHDQK
 KENAGFRL1SRGY1VGGSTMT1PQET1FAVAFSGFQKSDYV1SDIKQVYAGSLCAQSS
 V1VGLSSLRHLRLVSKVLP1ELGETPLVLHGQVYSGRNNHNTKLANNTQKGSDDWDS
 FAVEVGGSPVLDNLYR1YLSYSPYVKLVQSVNQKGFQVEADAPRI1FDASHLVN1PMG
 T1GHESAKPSGALLTLGYAVDARDHPHCLTSLTNGTSWSTFATNLSRQAF1FAEASGH
 K1LHGLDCFASGSCELRSSRYNACNGTYSF

mp_0445 494739 497579
 mp_7-Polymorphic Outer Membrane Protein
 NPLVSKKCLQMKSSVSWLFFSSSIPLFSSLSIVAAEVLTDSSNNSYDGSNGTFTTFVFSTT
 AARGTYSYLLSDVSFQNGALGIPLAGSGFLAAGDGLTFQGNHQHAKFAPINAGSSAGT
 TSSAADKNLLNDFNDFSLIS1SCPLLSSPTQGCALKSVSLGTNSGQ1TFNPNFSSD
 YGVNTKFNLLSGTSQSFASFRNQAFTRQGVGVVYATGTT1IENSPTGVSFQNLAKGS
 GGLYSTDNCISITDQNVIFDGNSAWEAAQOQGAALCTCTDKTVTLTGNKLLSTFNNTA
 TYEGGAISGLKVYSISAGGPTLFQSNISGSSAGQGGGAINASAGELASATSDGTFITFN
 QVINGSTSTSRNAIN1IDTAKVTSIRATQCSYIYPDI1PNPTGAASTDNLNLADANS
 IEYGGAVIFSGEKLSPTEKAIANVTSTIRQPAVLARGDLVLRDGVYTFKDLTQSPGS
 IEMDGGTTLSAKENLNLGLVNLSSLDGNKAAKLTAAKDNI1SLGSTLTIDTGES
 VYHNHLKASSTYPLLETLTACAGNT1LGLASTLTQEPETHYGVQGNWGLSWANATSS
 I65INWTRTGYISPEPSKRNLLPNSLWGNFDIRSINQFLITKSSDGPFFERLWLSGIA
 FFEVSDMSTTRHGFRHISGGYALGITATTPAEDQLTFAVQQLFARGDNIITKGNHDTYGS
 SLYFHHTTEGLDFANLFGKATRAVPLVSLSEI11PLSDFAKFSYLTNDNMKHTYTDYN
 IYSGWRNDAFCADLNASLPVVISVYLLEKVEPQKQV1YIAHQDGYVHEAARGAFN
 SELINVI1PGTVTFDGSSEKGTDTLMTLY1LDAYRRNPQQTSLASDANMAYGTN
 ARGQFVRAAHNLFQVNMHMLPQGFQAVE/RSSNNYNTLGSKFCF

PAE0446 497602 500415
mp-8 Polymorphic Outer Membrane Protein
IEEKKHLSMKIPLKLLISGTLVTPILLSIATYGADASLSPDTSFDGAGGSTFTPKSTAD
NCTNYVLTSGNVYINDAGKGTALTGCCFTETTGDLTFTGKGYSFSPFNTVDAGNSAGAAAS
TADKALFTVLTGSGNLSFYIAGPTVYVSGKSLTSSAGALNLDTNGTLTFSFGNVSEANNNG
AITTKTLISGNTSISFTSPNSAKLGGVYISAAASISGNTGGFLVPMNKGKGTGGGAL
FEASSISQNTSSLFFSGNTATDAAGKGAICYCEKTGETPTLTIISGNKSLTFEANSVSIV
AICAHGLDLSAAGPTFLSNRRNGCTAAGKGAIAIADSGSLSLNSAQNGDITFLGNTLT
TSAPTSTRNAYILVGSSEKATINLRAGQGSYFYDPDITNTGASVDLTINQDPSNSPLD
SGQTVFVSGEKLDAEKAADNFSTILKPLALASAGTGLANKGVLELVNPTGTQEGSTLL
DPGTKLKNGFYETSLTKLVLDLALBEGNKVSIEETAGANKITLTLSPVQDSSGNGE
FTINQAQTFPLVVFATASADSIYIDALLTSFVQVTEPEHYGYQGHWEATWADTSTAKSG
MTWVLTGYGNPERHRSVVPDSLWASFDITRTLQIMTSGANSYIQQRGLWASGTANF
KDKSGCTNQAFRSGKSYGYIVGSSADESFNISVAFQCLFGKDKDLFIVENTSHYASL
QALRFLFGLGPMPSGFTITMDKLPLILNLAQLSYFTKNDMDTRYTSYFPAEQGSWNT
GHEGFGSGLAYLPKEAPFFQGYFFFLKQVAYSRQNFKESGAEARFDDGLVNCNSI
VIGIRLEKIDEKKNKIFSLAYLIDVYRKNRPRSTSLMVYSASWITSLKNLARQAFLAS
GSHLTLSPHVELSGEAYLEGRSHAIYVNDVCGRLRYGF

NP_0447 500541 503351
mp_3-Polymorphic Outer Membrane Protein
KADPDTLYIKGSSLIWFLI LSSSLALPLSLFNSAPAAVVE INLGPTNSFGSGTPTTPPAQT
VVDPTITLYNLGTGDSY ITNAGSPALTALSCFKETITGLD FCHGCHYQLLQNLIDAGANTFTT
TAANFLPLSGSGSGYVLYLITTTNATATGTGALIKTCGAS IQCHYSCYGFQNFSDNDNGGALQ
SSGLLNLNPLPLFAFNKATKQAGLYCYGILITIMTLNAGSFSENTAANNAGLYITAEAS
LIGIHLIAFLF INNGVATFATQSGALCYCSGSGAPKPKVLTLLDKNLNLTIGNTAITSGGAI
LIDDLALGKSLKPLFKNNASLDTAAPLGGALADSGSSGLGALGIDITFEQNTVYKGAAS
QVITTFSSINLTGNTNAKIVGLRANSPNTIYYDFITGTLTADNLNLDALNGLPDLAGNPA
TITTFVSEIKELRPAEAEADNLKSLTQQPLTALVGLLPQVDTLVAKFSGSGSGSTLT
KATITLTETADKTLTINNLIVNVDLKKITKATLTKATQADQVTLTALGLSLVDSFGNYVD
NNTNVEVFSGLPLPLADINLITLIDLAADDLPLKNIILHWGLNNAWLQWEDFPAKGSAA
TWTFVYNDNPEPPPLTAVTLNKLISVYDVYILQGLVATYYRQETRTICWCSISNKK
VITLTHKIFRPLISGVVVTATLIDNLITAAFLPVLVYKDRHFLINKRASAYAAASL
QHILATLGLSLLYLGLGSLGSGVPLFAQVITVYFNTMTKTYTPAAGPKSSWYNDGC
IEALGALITPITALGTHYATFVFKVTAIVYHGGGPKFPTITLVRFDSGDLINSGVP
ITLTFEPEENKKAQYSLTLYVLYVDFKNNKXKALDILNITWKTITNLSRQAGIGRA

GIFYAFSPNLEVTNLSMEIRGSSRCYNADLWKFQF

[illegible]

CPr_0449 507231 505330
pmp_10_PMP_10 (Frame-shift with 0451)
EAYTGRFGCGGTSIFSNINVOGTTAGNGGASILAAGECGSLSAEAGDITFNGNAIVATTPQ
TLKRNSIDGGSTAKTLNLRASGHSIFYDPITANTAADSTDLNLNKAAGGKSTDYVSG
IVFSGEKLSEDEAK/ADNLTSTLKQPVTLTACNLVLRKGVTLDTKGFTQTAGSSVINDAG
TLTKASTEEVLTLSLTIPVDSLGEKGKVVIAASAASKNVSLAQPTLLLDNQNAGYENHL
GKTQDVFQVLSLGTATTTDVPVPTVATTHYIGYQDTGWTGWDPTDKTKTATLA
WTNTGYLNPNERQGPLVPSNLGWSFSDIAQIQVIERSLATLCSDRGFAAGVAVFLDK
KGKKERYRHKSGVAIGGAAGQCSSENLISAPFQCLPGSGDKDFVAKNHNTDTYAGAFY
HITECSSGIGCLLDLTPSWSHKPLVBEQGLYASHVNSNDLTKTYATPEVKGSSWGNNAF
MMLGASHSHPYELHCFDTPYSHYIKLNTYIRQDSFSEKTEGSRSPDNLFLNLSLPIGV
KEFKSDCNDFSYZLTSVLPDLIRNDPKCTTALVIGASWETYANNLRALQVVRAGSH
YAFSPMFVE/LQGFVFEVRGSRRIYNVDLKGKQF

CPn_0450 508121 507180
pmp-10-Polymorphic Outer Membrane Protein
SGFMKQSFSLWLVSLSTLCSTCVTAFAATEINGPSDFSGTNGTGYTPKNTTITGIDY
LTJTDITITLQGLDGAALTKGCFSDTTELSFAGKGYLSLFLNIKSSAEGAALSVTTDKNL
SLKFGTSITFLAAPSVVITPSPGKGAVCGGDLTFDNNGTILFKFDQEEENGAAGTAKNL
SLKNSTGSIIFPEAKNSVATPKKGGAICATGTVDTINPTAFLISNNIAEAAAGGAINSTGN
CTITGNTSLVFSFENSIVTATAGNGGALSGDADVTISGNQSVTFSGNQAVANGAIYAKKLT
LASGGGGVSPFLTI

```

Cpn_0451          508158      511058
pmp_10-MMP_10 (Frame-shift with 0451)
KTRQWRKIKILDSCFVIFNLIYLCFYDIANSSSLKXKSIITMKTISIWVLVSVSLAFSCHLQ
SLANEELLSPDDSFNKGIDSGATTFPTKTATNYSLTGQDVFFEYEGKQCTPLSDSCFKQITDNL
TLFLNGHSLTSGIFDAGTHAGGAASTANNKYLITSGFSLLSFDDSPSTVTYTOQGTLSG
AGGVNLENIRKLVVAGNEFTSDAGGAIGKASFLITLGTSGDALPNNSSSSKKGKALATTAGA
RANNTGVIVRFLSNIASTSGAIDDEGTSILSNKKFLYFEGNAAKTTGGAICNTKAGNSGP
ELISINNKTLFIASNVAETSGAIAHKKALLSGSGTFELRNNVSSNTPKGAISIDSG
ELLSAETGNITFVNTSLTTGSDTTPKRNAINIGSNKGTLLEAKRKNHTIFYFDPITSE
KTSDDSVLDKINSGAGALNPQGTILITSGETLTADELKVADELNKLSPQVPSVLSGGKLLQL
GGVTLLESTFSQNGASGLMGDSGTTLSSTAGSIITNLGIVNDSLGLKQFVSLTAGASN
KVIVSGKLNLIIDEGNIYESHMFHQDLFSLKTIYVDADVDTNVDISLSPVPADENSSE
YGFQGVQNVNMTDTATNTKEATATWTKTGVSPERKSALVNTLWGLFTDIRSLQQLV
EIGATGMEHKGROFWSMTNLFKHTGIDENRKGFRHSTGYVIGGSAATPKDDLTTFPACH
FLARDKDCI IAHNSRSTYGGTLTLPKHSHTLQPNLYLILGRAKFSESALKEFPREIPLD
VQVYSHSDNRNMYTSLSPESGGSWNECAAGIIGLDPVLVSNPHPLKFFFIPOMKVE
MVOVYSQNSFPESSSDRGFRSLGRLLNLSPVAGKALQFGIDGTYDLSGKFTVSDVRYNN
POSTATILVMSPDSSWKIRGNGLSRQAFLLRGSNNYNYNSNCELFGHYAMELRGSSNNYND
VGTKLRF

```

CPN_0452 511304 512860
 comp_12-Polymorphic Outer Membrane Protein (truncated)
 FLEDDTITLRNFLTCSALFLALPAAQVYVYLHESDGYGNAINKSKLEPKITYPCEGTSYI
 FLDDVIRISNWKHQDQEDAGYFINRSGDILFPMGNQYTFHNLMLTEGFGAAISNRVQDITTLT
 LSNFSYLSATPSAPLPPQQAQVYISLGSVMIEENSEVTFPCNGYSWSWGAAIYTPPYLLGSKA
 SRPFSVNLNSGNRIYFLVRDNVSGQYGAIATHNLTLLTRGPSCFENNHAHVDNMSGNATAI
 APGGSISISVKGSLDILFKNGTASQDGTIHNISHLQSGAQFKNLRAVSESGVYFYDPTIS
 GESHKITLDLVINAFEGKETYEGTISFSGSLDGLDHEVCAENLTSTIQLDVTLAGGTLTSLSD
 VYPTQLHQSFKQASSTLTMSPGSTITLCLSGDARVQNLHLITIEDTNPFPVIRIAEDKKTL
 VLEKLKVAFEAYWSVDFPQKFAETIPLLELLGSPFSDFLLSGTTLERTQVTTENDAVR
 FQWLSWEEFPPSLDKDRRTIPKTKTVTLWNPEITSP

mpn_0453 513156 516152
 amp_13 - Polymorphic Outer Membrane Protein
 NCVVLLLYFFYSLSLCIRIWHFYLVQMKTSIRKFLISITTLAPCFSTAPTEVIVMGSENF
 GSGSGKIFYPYITLSSDPRCTLCIFSGDYLINLNDIAISRTSSCSFNSRAGALQILKGQVF
 SFLNIRNSADGAIASSVITONPELCLPLSFGSGFSQMI FDCNCSLTSDDTSASNVIPHASAIV
 TLTPLTFTTNDSDILFIQYNRSAGFGAAIRGTSTI IENTKSKGLVFNMGNSITNNGGALTGSAA
 NLNINSAPVIFSTNATGIYGAAILTGTSMLTSGNLSGLVFNVNSSRSGQAIVANGNYIT
 SNNSSDLTFQNNTSAPQNSLPAPPTPPPAVTLPLLYGGAIFCTPPATPPPTPGVSLTIS
 SSSNSVDLTFLNIAISGCGALYGGKILISDNSKTI FLGNTAGKGGAIAI PESGELGSLSANQ
 ILFLNNKLSITSGTPTRNSIFHGKDAAKFATATGATQGYTLVIFYDPDITSDLSASAAATTV
 NPKASADGAYSGTIVFSGTILATEAATPATATSTLNQKLELGGTGLARNGATLNVNH
 TQDEKSVYIMDAGTLATTNNGANNINDGATILTNKLVLNLDLSDDGKAAV/VNQSNTGALT
 STSGLGLVKNSSGCLNNHGMFNKDLQQVILELKATSNVTITTFDLSLTNGYQQSPVGYQ
 TWFTFTITDTTHITLGNWKKITGLYHPHERLAPLIPSLNLANVILDRASVQSAASADEDPV
 KQGLSITGITNFNHNHTGDARSYRHMGGGVLINTYTRITPDAALSGLFGGLFTSKDYLL
 TGHCHSNVYFATVYNTISKLSGSSRFSGGTSRVATYSRNGEKVNTSYTLKPLGCSWSN
 GCHGLVEGLENLGLTSSRLINLNLQIIPVKAESVAYATHGIGQENTPEGRIFGHGHLNLVA
 VPVGVPFGKNLSNHRDFPYTIVVAYAPVYRHNPDGCTTLPLINGATWTSIGNNLTSTLLV
 NASHSTSDNSVDLELFGCGCDIRRTSROYTLIDGSKLRF

Pn_0454 516179 519115
 Pmp144-Polymorphic Outer Membrane Protein
 MPELPKRSQPELLACLRASACAFARLRGHEVPPITNGEGEILLTSDPVCNFLGASF
 VQVQINSGNLSGKGLILSTPTGQCAFTNCHYALLSAETILTKFNFSINFTNGCNLTGL
 CLTYGKDLVPSIKDLITFNKRVASPVATGATATITVTTGALQALQTLDTLVTENI
 QQIFVPGFNILNFGSAITGSSPTAVVKFINNTATSFSSHFNTSSGGQIVYGGSLFENIS
 CILPTANGLYCNESKVSTPSSTPYALGSGGALCIPTGTFLKKNQKQITPYNKTPTNDG
 IYAEYTNIVNQLALLLDINTAAPNKAIAICAVYLNITGGPPIEFSRNKPKGAIATIGP
 DQVPAKQKTLTLAGLFDLAFQGNMLNTYGLRNAITVEAGIEIVLSAGLRLVYF
 FTHGLPTTPGSKEDITINWAKYGVVPTPGLSTELLIPANTMTITLITVKAAGE
 KTLIAVAVVLEGFADGKQITKATITLPTTGAPAAVDITIGLKFADNPSLPED
 VQSAINATPTPTDGLALDEHATDLYIMVGLQVAPILAVFKNITPTKINFGG
 ATPTCHYVQGGKWTYWRFLIPATVYVGGSPSANTIYAVNNITLVRVYITLIE
 YCEITPCLNMLKSLNGLADSDLDGMLLTHGLGLITAKALGAVETPKPKHAKYR

FLPLMKHTRPHLKIIRKFLPLVGMVTRPPVPEDHKILLIKTEPLHIRTVFARVQDILL
PQGLRHITAAADILEPTTQCSGDIYEFYGTSEPAIDRIPLFEFTLEPYKHSFFFYRMDLQ
TLESPQEVFVIFVIESIPEGEDQAMFISKQGELELSDQSWIKPRIISPSDERHAREIQKH
IEDQPCFPFLKAMETDHTISQCVLGRYFSPASLKMFLSNYSRYLQHYFOIPSPSTSG
EFFSNRDSFLDLQYFAGISVFWADLESKRLLQYIKRRNKDVGMFVPHKQAEQFASQYFI
GTHGSCILAGDVEFLRELLTOMHTLSQOFTIPEFPPTPLAILTGGGSGAMELANRVAT
ELSLGSCNLSLDITNAYVEAKMSYAPDILLERQADPHVLDLAVFVIGMGTDFELLLEL
ISLKTOKKALVPVFLIGPVDVWKSKITALYNSNHAVGTIRGSEWVHNCFLCSSLAKAGIA
LQVQVHITLILHIFHVFQGTG

CpN_0491 574595 573336
CT389 hypothetical protein
ILSSLYTVFTMTAFHSCYSWFCWFLFVLVFGGIAGGEPLCPDCKYETKSVLRSDQLP
DHLWVYENDCYLTGVVQSLDMHFLDSRTQVVEIKNRAYLFLSPVDSLSSEAINTFVRDL
FFICAVEICERPYGECITRASAERPLLPKKEITLQMPFCGKEGVWLPQNTILFLSPLADP
RQVINSAGIRFNEKVGVRVGTIFGGDFILLRLFDVSRFHVDCDFGIGQGVFSVFLDLH
PESCMYNSDFVAGLWSGADKWSFRFLRWLHSSHLGDEFILTHPNFPRFLNSDEGVDFL
ISFRYTPQIRLYGGCGYIVRSRLTFFPERFFYCEWGAELRPFGLREGNLHAGTIFAMHKL
WEQKFGLEDQSYILGMEWAKFQIEGRKIRAVLEYHQGFSKEGQFIREPCNYGFRITYGF

CpN_0492 574643 574804
No robust homolog present in Genebank/EMBL as of 11/7/98
LFLSLFIPICEERNSSQTYKHLHVESACFLLESPLKIHWSPPYGFPPFYRDLKL

CpN_0493 575142 574955
No robust homolog present in Genebank/EMBL as of 11/7/98
SKTEGSHSKTSKGFVGRFVQWIRFTTGRGSKRSPSSPSTPHYRLATYTRSPKQSGVE
RKQEDAETSFIETPKGILKKPGNKDPKGKHHVWKDS

CpN_0494 575370 575146
No robust homolog present in Genebank/EMBL as of 11/7/98
VLMIRVNYGSGYGRNPSPEDGKKVPLSGNSRLHRRGGIRRKHSASVGTSGSKTGKA
SLKETMVLASEIAHFH

CpN_0495 575507 576793
aspC-Aspartate Aminotransferase
RRLLKNQKMAIQKAGAFRLCPSESRPLVEHAMRRNPFLSKLPQYLFSEISKLAQFRK
ENPEISVIDLSIGDTTQPLCRSITQAIKEFCVSOEQETVYRGYGPETGLEKLRITKASEV
YENRISPEEIPISDGAKPDIIRLSFFSGSEKTLGLQDPVYPAYRIAHTIGIRDIPLAC
RKETGFIPELBNQSLDILCLCYNNPTGTVLTFQQLALVNYANQHTVLIFDAAYSFAF
VSDSLPKSIFEIPEAKYCAIEINSFSAKSLGFTOMRLAWNVPKELYDNNEPMINDWKL
LFAITTFNGASLLMQEAGYGLDFFTPFKSLYLTAQAQLKKSLETAGFSVPEGNDHAPYL
WVLEPEGISDEEAFDFLLHQYHIAVTPGHGFGSCGQGVFRFSALTQPONIALACDRCTA
SLKETMVL

CpN_0496 576751 577812
CT391 hypothetical protein
PPLAFTFRKNDGSCMTILRKLQYLVFFSLFCFSFIYVATCGSQDPSVSSPKIAIFLSFPH
PLLEDCSKSCIEITLKDIFENLPIEIVLNAEDSIVKARKIARSLHTDKNVAIVTLGTIATK
VMSHETQKPVYIAVDPRESLTLPKNTMNIQVNDLTDINOYCFATQAVATNAQSIVYL
KPSPEPSPDLQKEIVKKLHAGSIEIVISITSTFKTRIRQAIDKRPASFIPLSPLSHK
EGTAFLOELKEKIPITDITDTSLSIEGACIACSVDYKSGQIAKIVHLLHNMHDVSL
RKITAQRLSPTTTTFNEDIKYLKIKLHKTERNQFLSFKSKLEKSEKGNVAVS

CpN_0497 578107 577820
CT388 hypothetical protein
IFQRVLLDSDWILEVKTVPKAKENKIVGFGQGLKVRVTEPPEKKGANDAVISLLAKALS
LPRKRDVTLIAGETSRKKKFLPNVRQDIIFSLHIDV

CpN_0498 579062 578085
No robust homolog present in Genebank/EMBL as of 11/7/98
YCRRLRAPFMNRKARWVVALFAMTALISVGCPCWQSAKSRCSIDKYIPVNRLLLEVCGL
PEAENVEDLIESSSAWVLTPPEERFSGELVSIQCVKDEHAFYNDLSLHMTQAVPSYATY
DCAVVFGGPLPALRQLRDLFLVREMQRGVRFKIVFLCGERGRYQSEIEQEHFDSRYNPF
PTENWESGNPVPSPSEEEIAKFWMMQLLPRAWRDSTSGVRVTFLLAKPEGGNRVANRK
DTLLFRSYQEAFFPGRVLFVSSQPIGLDACRVQGFPGKESYDLAGPGFAQGVLYKHWP
RIGELTLAEWLKETNGCINISEGCPG

CpN_0499 580404 579205
No robust homolog present in Genebank/EMBL as of 11/7/98
LSVYLLIFYPCNCTMSSVNOSSGTPNPEEVSPESTEENKVVSSDEAQATHAVALPIV
TQLSLPEGVGTSSBETASNPRVDIEAESSRVAQDISSLSVERVGLLDDLKGAGSLF
TSFQSELKNCLPAWKSSTRLETRGAGDNADLARLEFRSDYEAVLGHANQPHGKAHLIT
SKLTDVHHKLQGLSREDLSLAFDNNDRVLEHLGSLGLDVAEGNWSLSCERICPLRLVLA
DSMLVQIKKVNLPTEELRLTQGTTESSSDPRVESLSCCERLLNELRLWANFVGFISS
CYDNIFFVLMWIVRRINILLPGLGCLPFHNPDAEQSDQSSSGERSTRERLRSRDLSEE
EMIVRAEGESIHPESPHGDGRNQPSRGDQSDSSEETEL

CpN_0500 580647 582362
proS-Prolyl tRNA Synthetase
QPHSMKTSQQLFYKTSKNANKSAVLSNELLEKAGYLFVKVSGVYTYTPLLWRVSKMMNI
IREELNAIGQELLPLHNAELWQHTGRWEAFTSEGLLYTLKDRGKSHCLAPTHEEVI
CSFVAQWLSKRLQLPLHLQYIATKFRDEIRPRFGLIRRELLMEDSYTFSDSPCEMNEQY
EKLSAYSKIIFORLGLAVIVTADGGKIGKGEFVCLSGEDTICVSGSYGANIEAA
VSIPPOHAYDREFLPVEEVATPGITTEALANFFSIFLHKILKTLVWLKSVSEKFI
QMRGDRQVNLVHVASKNADIALASDEEIERVLTKQKOFIGPLNCPIDFPADETTSPMT
NFCVCAQNAKDKHYVMVNRDRLLPQYGDFLLAEGGTCPENPGHPYRYYQIEVAHIFN
LWTFYTDSPFVNFQDEHQQTQQCQWCTGYIGVGRTLAACVEQLADDRGIVWPKALAPFSI
TIAPHQDQVYQELAETIYHELOQGYEPLLDLDRDLRGLKDSDLGITYPYLILGKSY
QAGYHIEISRGKGYTVSPAEPTWQCNHLA

CpN_0501 582121 583650
hcrA-HTH Transcriptional Repressor
ILLTFYXGPPIMLGVTVLVGLFMARIKVSKPDKILDLIFATTELYLKTQCPVGSRTKL
KSPCLDLATIRNYFAELGALGLKKNITXGRIPDGLARHYVDHQECCFCAEISAPI
FDKICQLPEENILKIKQATELLGLDLPTFFGPPFNDSVTNIQITQVOKRAVT
ILTFEYQITPTDNLWLEACDTLAKRIEKTIONYPKPLPTNEELKKEHLJCMLYNEV
VWVILTRYNMFDEEDLYQTMKLLKYCAFKDPEVLAUGLFLNRYKMLRLLHGMKQ
KATAFIKKILKILHATINPCSVITIPYMYNRSPLGALGILGILNRYKFLPLKLFAN
KINETIQYFYKPKIITPRPLTNCKLNFPILRTEYGTIKLPLPPTTL

CpN_0502 583641 584201
grpE-HSP-70 Co-chaperone
GDVMTDTPPENEEQHESN/VNENE/EHLQOEITVTLKTELKEKNKYLMALAESENKRL
OKERQELMIALENTLIDFLNPISGMEKALGFATQMGDDVKNWALGFNMLNQGKIFEE
KGIIEYSSIQGFNFPLHEAVQTEETSEVPECTILEEFAKGYKIGERPIRVAKVKVAKAP
TPKENKE

CpN_0503 584225 586213
CT421.1 hypothetical protein
PAKRUQAVTNFEKTLGSTRFFIGRKYSEVASEIQTVPITVTSKSGDAVFEVDGQYTPPE
IGAQILMMKMETAEAVLGTEVTEAVITVPAYFNDSQRASTKDAGRIAGLDVKRIIPEPTA
AALAYGIDKVGDKKIAVFDLGGGTDFDISILEIGDGVFEVLSTNGDTLLGGDDFDEVIKW
MIEEFKQKEGIDLSKDNMALQRLKDAEAKAIELSGVSSTEINQFPTMDAQGPKHALT
LTRAQFEKLAASLIERTKSPCKIKALSADKLSAKIDVLLVCGMSRMPAVQETVKELFGK
EPNKGVPDEWVAIGAAICGGVGLGGEVVDVLLLDVPLSLGIETLGGVMTTLVERNTTIP
TQKQIFSTAADNQPAVTIVVLOGGERMAKDNKEIGRFDLTDIPAPRGHPQIEVSDID
ANGIFHVSAKDVASGKEQKIRIEASSGLQDEIQRMVRDAENKKEEDKKREASDAKNEA
DSMIFRAEAKIKDYKEQIPETLVKEIEERIEVNRNALKDDAPIEKIEKTEVDELKHMOKI
GESMOSQASAAASAAANAKGPNINTEDLKHSFSTKPPSNNGSSSEHIEEADVEIDN
DDK

CpN_0504 586418 588514
vacB-ribonuclease family
ATQPTSETTGFLVQCPKLTGGAQLLKKPKRKGPRRTYKSLKIFIPGTLFVHARKGFGV
SPDNPEEYFPDIFVFPARDLRGALGDHIVSVLVPYPRDQKLGKGTISEVLARGKTTLVGT
ITSVSPSALATYSMSGSSQLIPVELLPGRYTKIGDIRLLSTPPWVDKPOEGASPAIQM
LEFIGHTAKADFOAIQAEYNLAEEFPPEVIEEASLFSOKHITQVLSRKLDRLLCFT
IDSSTARDFDDAISLTVDHNNYILGVHIAVDSHVYVTPHSHLDKEAAKRCNSTYFPGKVI
PMLPSALSDMLCSLKNPVDRLAVSVFTFTKSGHLSDDYQIFRSVIRSKYMTIENVDNI
EKHSHSKILNEMATLSKKFSDIREERGCTIRFVLPSVTMSDNLQEPVALIENHQTFS
HKLIEEFLMKANEVVAYHISHQGVSLPFRSHEPPENDNLLAQELAKNMGFDITFTPTQE
PDVQYLTQTSAGHPLEQVLSQFVRSMTASYSYENKHYGLKLDYTHFTSPRIYID
LIVHLLFNPLSIDQTHLEIIVRACSTKERVSAKAENSFENLKKTRFINKFLQEQPKTTY
HAYIITANHEGLSFVTFEFCHEGFIAAELPKESLKNALPESIPDKMKPGASIKVTID
SVNLLQKIVWSIATTTEDPKKIKKTPSKKKGTKKRA

CpN_0505 588471 589106
*3-methyladenine DNA glycosylase
RKRLLRKERRKKEPRNVLQEHFLESDVITLQAQLLGHKLITTHEGLITSGYIVETEAYR
GPDDKACHAYNYKTRONRAMYLKGGSAIYRCYGMHLLNVVTGPEDIIPHAVLIRAILP
DQDKKELQRRQWDRKPHLLTNGPGKVCQALGISLENRRORLNTPALYISEKISGTLT
ATARIIGIDYAQYRDVPWRFLLSPEDSGKVL

CpN_0506 589055 589840
CT421 hypothetical protein
CPMEISIPRRFGKSLILNKLKYSKETNAHLISCRRIMKKYFITGLVILLPLAITIAI
VTMIMNLFQPPFVGLASEFFKFSFYTKHRAVLLKFLVLIILLFGLFFATVLLGFLTRIMI
KSLLSIYDKLHRIPIIKTVYKAAQVMTTIFGSKSGSFQVWVMPFANVOCIGLVA
GDAPTCTCGKEKEDDPLTVFIEPTNPNTSGFLTLFRKSDIVFLDMKIEDAFKYIISCGV
LSTPMACPSPLPDELHQDGS

CpN_0507 589898 590122
CT421.1 hypothetical protein
STPYQPLSGEIKKFNIELFMTRMSQARRAKSPKKRKPYAIVHPAPAPRIVYKLHT
NALSTSDSIFIPKIG

CpN_0508 590133 590300
CT421.2 hypothetical protein
SRIMSRHSYGVKGVTKRNVLRKRFERFEVLRKLGRWNDSATKAVTGLPKPTILK

CpN_0509 590299 590808
(predicted Metalloenzyme)
NKVFVLYGNFIRVTOEKIKIHVSNEQTCIPIHLVSVSEKLVLTLEHLKLVTTNEIFYFLE
DKALAEFLYGFADPSLTDITLPIADPGDPAHYVLGEAFISPOAALRFLENTSQD
IYEISRYLVHSILHMLGYDDTSSEKRRMRVKNQIILCMLRKKHALITA

CpN_0510 590804 591973
tlyC-CBS Domains (Hemolysin homolog)
QLNMHILLALIFCILLFLAFLGTQPSCHGSSKFLKTLNRQFFKDKGREYPPFAPSAPTILA
TLICLVGALGTGLYTLPLPKTAHKLDFWPLVLSLALIAVGLFPWISTKVPKETTALH
RFLASVGLKFLPLQLLYRRRNQVRSSTFSQSLSEALSFAFNILIREVMIPKVIDI
ALPEETTLQEAALVLSSEEGYSRVPVYKKNLDNITGILLVKDLLLYTSSHLDQPSISSA
KPPFYAPEIKASSLLQEFQKRRHLAIIVNEYGTGFIATMEDIIEEIGETADEHDQ
ENTPYKIGSSWIVDGRMNIADAEYFNLIKIDHENSYDTLGGHVPHKVGAVPQKGRMIRH
ENFDIEITCTERNVGLKITPRKRKNIS

CpN_0511 592141 592488
rsbQ-Sigma Regulatory Factor
MSDIQKEHGSTFIHLHGKLDGSSPEVQENICQSLAAGSKNIILCAHLDYMSSAGIR
VLLQSYHQVQGHSGKIVLTVFKTIEQTLVYTGFLSYFKIFNTVDEAIQTLNKDGD

CpN_0512 592538 594412
CT425 hypothetical protein
SLPLTMRSSVCYNPSIARAGQISTWKFYLSLATPLPAGTKCKFDLAGSGKPTDWEAPAT
DLSQTRNVYIAEMPEGEIEATAIPVKDNVPVQFEFTLYELQVGETLTIVMGASPNHQ
VDDAGNAQPLAQRKPPYLYIDPTGEVNYDEPDVFSMDIPGNVLKLEIPTPSYVKNK
RFDITVRFDEFGNLTNPSPEETRIELCYEHLPEZLNWOLFIPETGFVILPNLYFNEPCI
YRIQLKNLSQOEIISAPIKCFADSAPHLMWGLLHGESERVDSSEENIETCMRYFRDDRAL
NFYACGFEHGDNLGPDWKLINQTVLDFNEEDPFTLSCQPYSGEPHLEGVRIILHTEK
TSHSKHKEKHYKILAKLYKSTVNIDMISIPHTASKEHGFDFENFEPFVEEYINAW
GISTTAAALINPFIQKDSSEDPRTVIEGLKPLHRCFGVAGGLDDRGYKDYFDSPOVQ
YNYGITAICNKHTRISALALFAHICATTPPVLFSNITCAPMSLELTKSGKPLGNV
NIRLUXHVATALLKTVIIRNFVITTFPFDNRILDEYDDMVPLSVTLKDPHKGKAPF
VYLYPLPGAINAMAWCPFIWVDIN

CpN_0513 594515 595751
CT426 oxidoreductase
VHICLAFGLKLVKKNLAAKKHVTQGVTVLQHLIERIVHQSIVQMTCILPQPKTSPLYS
IFKILDAQEPLESDALHLLLNKILQVTLWIFADQVPPQPVGDYVYVSTLYLPTNF
CIPKXFPQYAKYKIPKMLYVVDLILQVLTQHTPTITEYHIVGCGPQCNLQYQSLP

CPn_0546 n312d3 n30912
r121-L21 Ribosomal Protein
LSKQRLTSLSERIRKKLMPEYAVITQTSGKQYQVRGGDIVDVLLEGFASDKEVIFQDVL
FVFDGTKASLGSPTIANAQVKAEYLSHVKGEEKVAYKYKKRKNYHRKKGHRQKYLVRKIR
ELL

[illegible]

CPN_0548 633234 632191
cysJ-Sulfite Reductase
KMVLEQKFKAAQVPLRLNLLSCDSINDSDPIYRMVFDNSDNTTISYKVGDAVLGPENIS
KEYSEVHVLQQLGYSPTLLRKKSTSEKVSQAQKIQGVVLDLKPARKINSVFDPKDKPITLL
YDAIQEYRPQIPIELFAESYVPLLPFRFYSIASSDPLKSTELLVHKVSYPGKYQKRFVG
CSSFLCSLQVNDASAYIFVQPTHTLTSTQTEKPLVMIGAGTGIAPYKFALEERLFNKD
PGNNLLFLGFERKEKNVFYIEFHNAAEEGKKLKLFLAFSRERHQVQVYQDILLRKQKDEVR
KAYEGCGFFVRCGRVKLIEVHKHAAEEILGKLTASLRKRDQVYVDVY

CPn_0549 633662 633255
rs10-S10 Ribosomal Protein
PQDVQHQPWNQHSLLRFLKKFKRLLRSKGCМКQKQKIRIRLKGFDQGGLDRSTADIVE
TKAKTRGAVRVGPIPLPTKREVVTVLRSPhVDKKSREQFEIRTHKRLVILDPGTGIKIDAL
KMLALPAGVDITKIAA

CpN_0550 635688 633580
fuaS-Elongation Factor G
LNYNGNNKFMNSNQEFDLASAIRNIGIMAHIDAGKTTTTERILFYAGRTHKIGEVHEGGATM
DMAWQEQRGITTSAATTVFWLGAJNITIDPGHVDITIEVERSLRNLGDVAVFVAVS
GVEPQSTVWRQADQGVGPRIAFVNMKDRMDAGYFAAVESMKEKIGANAFVCPGIGS
QFVGMVPLISOKALYFLDDTLGAKWEEKIESIDLKERCALRANLEELATIDESNEAFM
MKVLEDDITDEIHQVMRKGVINENINPVLGCTAFKNKGQQLLNVIKWLPSPLDRG
NIRGINLLTDQEISLEPRRDGPAALAAFLVITDEPYVGRITFRIRYSGTLKKGSAILNSTK
DKKERISRLLEMANERTDRDEFTVDIGACVGLKFSVTGIDTLCDDNQEIVLERIEFPDP
VIDMAIEPKSKGDKREKLAQALSSLSEEDPTFRVSTNEQYTGQTISGSGELHLDILDRMT
REFKVEANNGKQPVSYKETITVSGNSETKYVKGSGGRQYAHVCLEIEPNPEPKGKNEVS
KIVGGVPIKEYIPAVIKGIEEGLNTGLVGLADVVKVSIVFGSVIVDSSSEMAFKICGS
MAVDCAKRAKPVILEIPMKAVITIPEDHLDNIGDLNLRKSGKILGQESSRSGMAQVNAEV
PLSEMFGGTISLRSLSGSRATSMPEAFKPVQKIQEEIVRK

CPn_0551 636174 635698
 rs7-S7 Ribosomal Protein
 MYMSRRHSIAEKRDIPGDPYIGSVILEKFINVKVMHCKKSVA RKIVVSALERFGKKLNLEN
 VLGGFGAEALNAKPILEVRSRVGGATYQVFVEASERRNCLAMQWIKHARSQKGSME
 VGLATEIDCFNPKQGATIKKREDTHRMAEANKAF AHYK

CPN_0552 636698 636219
rs12-S12 Ribosomal Protein
IQAGYVPSSSENKPLPTKRALLYISMLVVVRLKREYMQFTINQLIRKRRKSSLARKKSPA
LQKCPQKRGVCLQVKTTPKPKNSALRKAVWVLSLNGQVEIATYIGGEGHNLQHSIVLIQ
GGRVKLDPGVYRIIVRGTLDCAAVKNRKSRSRYGAKRPK

CpN_0553 637753 636812
No robust homolog present in Genebank/EMBL as of 11/7/98
CGMMRVRLFLFLIIFILGRVAFLPRASEFSWETSTKCLPVLGPIPIITLTNDEVAQCG
LQGITITSTNNAKIKEIFLYIKERFPBASIFSRKCPENLQSGHLSDLGLITNMRGETYA
EGMANGKENGPAKQPKDRLRLVLRCPQDPTLLYSEKAEAKGITETNCLQSGDATTLDQGL
ILYKIDSIEKFLKETKRKNNTLVDLDCPSQVTTFFGRFWSLLNVLQVLFILNEDSAGIL
IPDLAQAQTLLSHTVPLLFYITNDSIHIEQGGSESSFTYNQDLTEPILGLFGYINRSGM
EYCFCAQQLSSLGET

CPn_0554 637806 638141
 CT4740 hypothetical protein
 VFSLYLLLCIILVYVRFMYECKSRMASPTPGQLHLQKQKVESKAYDYSRSLAMIATALLFFI
 VALILSGLSLLPQVFLPFGSGAYFIIGSLFALFIALGILLINVCVDLQYLTSS

Cpn_0555 638298 640241
tsp-Tail-Specific Protease
MFVFMKGLVRLCVLLSLPNVLFSQDLLEEGIKMKMDGLIEYHVDAQVSTDILSRSL
SYIQGFDPHKSYLLEASQYEVAVFLQSPDTEKKRLIKYKAGFAIYRNQINLHESILRAQW
RNWVKNPKVELLEASQYISKQPMQWSKSLDEVQKQRAILLSYLSHLGAGASSRYEG
KEQLQAALRCQLTENHNVYGLIDHNGVAMDREAEAYQFHRVYKALAHSDAHTAIFYSK
DEALARMILQLEKGMCGIGVVLKEDIGGVVIREIPGGPAAGSGDLQGLDIIYVRDGDGI
HLSPFRGLDCLCGHGGSTVVLIDHRSGETDITALLRREIKLLEDRRNVSYEPYGDVGIGK
VTLHSPFYEGNCNQSQDLRLALQKEKNILGLVLDIRENTQGVLSQAIYKSLFPMNT
VWVSRYADGTMKCYRTVSPKKFYDGLVLAIVYSKSSASAEIVAOVLQDGVALVWDGEQ
YVGKQIQTHTITGASQDDCFVTKGVYVQYISPGKSTQLQGVKSDILPSLYAEADRLGR
FLEHFLPADCCNLVLDPLDLDLTQTRPWFQKYLEINLQKQETLWREMLPOLTKNSEQR
SENSENFAQLSLKSSKETDLSYGSNDLOOEPINLMDKMLLOOCRK

CPN_0556 640921 640325
CPN-15kDa Cysteine-Rich Protein
ENCMSSNLHVPQGTQGAAPESVLNIVEEIAASGSVTAGLQAITSSQGMVNLIGWAKT
KFIQIPRESLKFQSRACQITLLVLGILLVVAGLACMFIPHSQLGANAPFLWIIPAAIGLIK
LLVTSLSGFDCAETSEKLVLMFQKAGVLEDQLDDGLNNNSKIKGHVKTETCINTSRATFVPL
NDGTSRTPVLPVLSKIKARV

PEn_0557 612870 641194
 tmcB t ukD1 ynf40c1-rich omp
 LPEKLEKLRIRIVTVLQAPMCAQCFAGFICAAVAESLITKIVAGAEKPAIVMTAFYKVR
 LVRRKQPEVQKIRGAPADKRFVYCEBFRQVPAQAEQVYGRLYSVKYNLQ NVICQIS
 VPEYATWGSIVLILLALGRKK YDVVITQQLPEVAFVSSDEITPTTIDKLKVMIDRL
 IAGCKQITWVWKTKLKKRITATVYCAAPELRRVTRIGQPAICIKQKIVTACILKQFVC
 VLVVNTVILARNVTVNVPDVGACACACQGRVLSFNLGDMDRPGDKKVTVTVPQGRG
 ITNVAVTITVGRHRCNANTVTVITV YVVNLSADNYSYCKIPVEY NVINIDILYLI
 LNVYLQPLPRLVYVLAIRKILIPYANIKPEKGEGLQKLVYKAGVIVKRTINAV
 IESNCKITITSCALTITVGRILAAHTHVLVLPNDICVGNITNIVITVNRNIAELTIN
 LKILKPELQGLAIPKPKITNTVYDALPVLAKKEVVEFVTLKGLADDAIGAI

tsf-Viral-Specific Protease
 PFMVKKLVRLCVVLLSLNPLVFFSSDLLREEGIKKMMDKLIEYHVDQAQEVSTDILSRSL
 SYIQSFDPKPKSYLSNGEAVFLQSPFTKRLKLNKYKAGNFAYIRNINQLHLSILRARQ
 RNVEYHKNPKFLVEASSYQISKQPMQWSKDEVKQQRALLSYLSLHLGASSRYSYG
 KEEQLAALCLRQTEHNENYVLGINDHGVAMDREAYQFHRVVKALHSLDAHTAYFSK
 DEALAMRIQLEKMGCGIGVLLKIDGVVIREIIPGGPAAKSLDQGLDITIVYVDGKDIE
 LHSFPGVLVDLQGHGCHSTVLKDIHRSDEGTLALRKEIKLEDRLDVVSYPYEGDVGVIK

SpH_0541 621737 628003
 Solute binding protein (-yebL-Synechocystis Adhesin Homolog)
 NNRSSYQTAFVMHKVIVIFLTLTSLKSYGNDVIDKPHVLVSTAPYKFLVEQIAEETCFV
 ALATLNHYDPTHTLTPSPQIKELRGDGLWFRIGEAEPCTERNLTQQQVDLSQNVSLIQ
 PCPCNQTHTNYDTLHPKQLNKQVETIVTLTSLKPHQATLVQSNQKLLLDNLDQLE
 ISLITLTSKAKQRLVLSHGAFAGYFCRDYVNSDPTIEKSSVPEPSKPDVARVFRDIEQYKI
 GSVLLEYSGRRSSALADRFMHMTNVLDPYAEANVLNLKTIATTFSSL

CPn_0542 628000 628737
ABC Transporter ATPase
MTMIRILAEGLAFRYSGKPGNIHDSVFSVYDGDFIGIIGPNGGKSTLTMILGLLTPT
GAVSLDTFPFSAGQKTHSMIGVWHPQFSDPFCPIISVDKDWLSGRLSQLSWHGKYKKKDF
CSYDHALDVLGLSDHHHGHGFWHLGGQIQRWLRLARALASYEILILDEPTINDPDNQQR
LSILKKNLRTCTILMVTHDLHHTTIFNFKVFMNKTLSLADTSTLTQDFCSHPKYNQE
SCSPH

Pn_0543 628710 629603
 (Metal Transport Protein)
 KSGIFMLSSGLIRDSPFLLLLTPTLALAGASVAGVGMGTYIVVKRIVSISGSIASHAILGG
 IGLTLNIQYKHLHSPFMYGAVTALFCALICGKTHLKYQEREDSLAMIMSWOAGIIT
 ISRLTPEINLEINFLNGLIIVWTPSDLSLYGIFDLVLVGLIVVLCHTRFLALCLFDERYTA
 NHCSQWLMYFLLLVLTALITIVMLTVYGTITLMSMLVLPAVIAACRFYSKMYIRMFISVL
 NLCSGSGICIACYCLDFPVGPTISLMLGTYTASLCVKKKRYNPTSPSPVSEINTNV

CP1_0544 b25948 b25925
hbp-GTP binding protein
GKVFVFVTHKGSFFFLCKKMYIMFVQITLLELRACKGNGVVAWRKEKYLPGGGYVGGN
NRNKKVIVLEATTQVYFKAYRNITFLFKAEQIQGATINRTGRCKDLVSGVPTQLLRD
NEMVEILHFDHPTVDGERLHLYVQKIKKZKNCITFKFTGVNRAPTKPAFKKRLQVLEKELK
ADGLGVLEHMAKKTLLPNTLQIAHTEVRYNAYPTLAPGLGGLVKCKRQLYOKPMIADTP
LIVLHAKNKIKLIDLRILNQLLIDLLFLVFDVKREPSDEEDLELHLELHCHOPDFEK
DMMLVALNIDDLLEDFDLQLQFQKRIEPTVYELDGLNINIVGLVRFETRLAV

127 127 ribosomal protein
 TEAFEFPPVMAHKEKQXLSNPFRDQPKRLVYKVGWQPVTPQLVLPQRTPTFWHPAQNVG
 GRIQZPLFALVGIIVMKKINRTYDVAVEQL

LESS THAN FIFTENTH?

CPn_0558 64306 64301
omcA-9kDa-Cysteine-Rich Lipoprotein
KLMKKAVL[AAAMFCGVVSLGSCCRIVDQCFDPCAPSSCNPCEVIRKDS CGGNACGSY
VPSCSNPGCGSTECNCGSPGVKCTSPDGRCKO

CPN 0559 643700 643937
UNRECORDED COPY OF THE MESSAGE
FELARSVENTCYNTNL

CPn_0560 645666 644098
glxG-Glutamyl-tRNA Synthetase
RNSRFQGMKSLWSDKRRIMNWNVVRVRVAPSPGDPHVGTAYMALFNEIFAKRFGKGMIL
RIEDTDRTPQLEAEYAFISALRWQGLQWDEGPDVGPGYQRQSERATKIYQGVYETLLK
TDCAYKCFATRQDEAMENASTLGYRGQYDRRYRLISPEEVAEREAAGQPYITLRLVPL
SGCEVFEDYSKGRFVFWADVDQVLVKSDEGPTYHFANVIDHMLGITHVLRGEWLSS
TKPLLLYLEWPEWPFVLMHPLLNDPTGLTKSKRNPTSFYRDSGVYKEAFVNFVLL
MWSMEGDGEIYVSLERIETFPNIRKGSAGVFDQKLDMWNKHLYNHGESPCEKLKELQ
GMLNDEFFKLPLQCSRIITLAEFINLTSTFFSGFVREVELLPALNSPEKAAIILY
SYVYKLETKDQWTKTGLGYSKWLQAFNVHHKKAIPLYLYATGKKQGLPLFDSEIIL
GPKRARARLYAEKELGYPKKLAAVDKFMDFDEEATFDL

CPn_0561 646407 645871
 euo-CHLPS Euo Protein
 LMACEQHEGCVLEEREFEIDIKSDTKWVSITQAAKLHNVTQAIYVAIKQKLLKASKE
 TWAEIDIKDLEEYKRNRYSRKSLYQGLVFDNGKGCYSINQVAQLIGPVQKVYYATRT
 GTIRGERKGAAWIVHSEIERYKNEYLSKQAAKLLKGAEPKEHQAPNFPETPFESN

CpN_0562 648051 646918
 *CHLPS 43 kDa protein homolog_1
 NYVKYIMSTIARQYAAALDMMHPKPSIAMFSSSEQARTSWEKRRQHPYLYRLLLEIIGWVVK
 FLFLGLIFPLIFPLGLFGLQKIQCNFILLGAGGVIFRPICRDNLRLQQAARFASPSDFD
 VSVRRVCLVQYDEVFIWDLGRLPNAPKPRDWMLISNGNSDLEIYRTVYQKGDWIPRTAE
 ESQSNLILFNYPGVKMSVQNTTRNNVYKSGACVRYLRDEPAGPQARQIVAYGYSLGASV
 QEAELSKEIADGSDSGVWFVVKDRKGRSTGAVAKQFQISGLGVWLANTLTHWNINSEKRSKD
 LHCPELFIVYKGDQSGNLLGDGLFKKCTCFAPFLLDPKNLEECSGKKIIPVACTGLRHHDIL
 SDVYIKKQVAGHTHFDNFN

CPD_0563 650113 648293
 CDS ssDNA Exonuclease
 YQVHLLWDFSPKPGCGIKFMTNSDNASAGLLWAHPKEDPAFLGMIKEFHLPTTVAQIF
 ISRGFTIQEIKHFLYSHLSSYDVGPLFLDMSKAVERLLADRKHEHMYIGSDVDGMT
 GAVLLEFLRDLDIVHVSYPFLGAILQKHETSITLAKLKEEGLITITDVCGTAGTACVES
 DITRGIDIVITTDHMPGTGKIPHCVATLNPKLRDHTYPNRELTVGVGFKLARGVLNALI
 SRVSPKSSQSKGLKLLDLVLGTITDVGVLGVNRMVMRYGIKEIARGARPGNLKLCALC
 GVKESEVTSITDVLKIAPIKLSNIGRLDDPARGVENLLTDQDERDALIMELDNINRQRG
 IEAEVFDQVOETPLGNSNPEILKQAIVISSTAWHARVPTIISARLAKTSPVPVVIATISG
 IGGKSSARTIGSPLPLNVLKKCSLLSYGHHDFAGVINKEDKVEPKKFVHVNSSLK
 KQDILSKHLEIDAYADFDAIDYDLLASMELEFEPFGKNLMPIFYSKVRQVRYPKVLPGNHL
 KLLPQKNELEGVGLGRHADALKASWHYPLEIATYTPRLSQTSQSGSVIHLVRDRFIS
 SPERSFD
 E

CPa_0564 654359 650145
SecESecF-Protein Export Proteins SecD/SecF (fusion)
SSGDKKQVKRNFPAITICVFALALYYVLTCLYYAKPLDKKKIDGNEAEHIKISFTKQAQV
LRKIVIPRVSAILSHSLLRHGHIQHPAIPDIVSVFRKGGEDAEIGLNVHGPNVPKSA
RLKIVYSGREHDDHYIQVASSINTLVSDFSVFSVYSSENEQEMASSILQRYVSACTPFI
QKDCSCSYPSINWETAPKEQLLYAKNLSLSDGVEFSRSLSAFCQFSQSSNDRFLAFSLRS
LSHDAIAIDVQDKLLKSVYETLSQTACIRISGCPYIEGLRLDCSESSLFSLFSEYCPKE
RKETILTHSDLLAQTSLSKSEQRDLDSRLAEVQKLSKNTLVQVEDVNGNFSFOWMKD
PQKIIQLGGERLLQGI AEHTALTLRPAESCDLIPENFVFCRQPRESEAFGYIFSP
NDLCKHFHSGKSVYLLVKKGLRSIVAKYQGGGKLEQSFEDKLQNLNYCSFHTEIASWTLE
QVLEIRLQFLQFLVDVGGIVKIEGCAFALEVKDIQDRLATVNI EKNRQSDLVVRHEQ
YRHAKCSMDKQERLSPAPIYQNLFLNMKNLKRSGRGENILRLGIDFVGQRLLSLFKD
GHQKLTDMKEDILKVSDELCAIRNLKGVEISELIRREGDYIHLVSVGSSITSSEILGTSP
SPHFVNNRFSYSYASRYEQRFLDLVNYFTSQAGQTSPEEINTFASALNEEVDVPPSV
HEAITKLKSEGLFAFSPGCSETPTSLDITFSMAIGKDAEQKANPLVIFRNVNLDGASL
KDIRPEFAAEVGLVNLFSVKDITSPKKMAEKLSPTESFHTWTSAYCQEGISGTANGQYSAN
RGWRMAVITDGYVSSPILMNLVKHNASVGSKFTREHVSKLASDLKSGAMSFPVEVLESE
TISSDLGKKQCTQGITISACGLMALVILMSVYRFGGIVTASGAVLNNLLIWAALQYDLA
PLTSLGSLAGIVLGMMAVDANVLVFERIREEFLSLQSLKSKVEGYTKAFGAI FDSNLT
VLASALFGLFDTPG IKGAFALTILGI FSSMTALFTMKFFPFLMNNKTKQHTQLHMMNFV
IKHDFLGRCKKLKMAVSGSVFLVLCGLPWGAGVNSVLMGDFKGYGATFNPKEHGISDVA
QMRQVYVGLLQKLEAGLSRDRFITQGSSEKIYTFSDKALSYKADTSLSPKINDHEL
AVGLSLSETGLDFTSTETLNQFWSKVSSKSLKMPYQATIGLGLALAILLYLSRFREW
YAFSAVCAALIHDLTACAVLI IAHFLLKKIQLQIGALGALMTVGLYSNLNTLIIFDR
EDROANGLTIPMHVLVDNQKQTSRTVMTTATTSVLVLMLEIGGSSVFNFAFIMTIGIL
CTLSLLYIAPPLLLFVMEKRNKS

CPn_0565 655741 654533
CT44y hypothetical protein
NKLGFCLFICFVNISAILPDPSSFLLIKIKRNSKRMRLSMKFPFRISISDLIPTQMVIWWRGG
GVNYSVPAQNALPKKILGVLGACFLGLALGCAAFVAAGVCOTIFPCIGLMIGLPGVLGFAFY
LQYSGKVSFRFPLFRETUKCFKPINWLGSLSLQSNKTLRPGCYHGPCGVQVEICBGSQ
EIVTLFQKKSDDRNSTSLFIQEMDQIALRQGIKESLSLRKTAIDFSPVSVSLSEIQORE
QYQLDLPKVLSWSSDQASDRTHPKSAIVVNISDAAEQPGRCYIDATYKFTTFLVDQIGD
HIVTKGTTIYVLTPLIGVPDALPKEEQENKLLSQAAFLYSAEQVAKRMREEKQSDTRIK
FIDFTDPTKTLGYPSPHSGSTHSPVTPSLGSGFQOEQSYTFA

Cfrn_0566 656019 656890
zao3 family
EPLVTALINDPVDILSLATNNNAESYFPISLQPLPNHVAIIMDGNRRWKKHREECCHTHTS
CHYYAKVLPHNINLAVILDLQIKVLTLTPTOTENFGPKKEIQIFNFIPTYTDQLKPLYLE
HNEHLDELIGDLKDLKPLKGIQTKLNHVRMTAIDPSLELVLANVYGGKDELVARFKLHVD
LHKEIGDGLGELGELISYLLSLTGLTDPDLLIRTYGEMRVSNFLWQIAYLEYITDTLW
IDPTGHLFEALINVYQQRSLRQK

Clu_0567 656894 657817
GluA Phosphatidyle Cytidylyl transferase

VLNSNKFKSKTQAYGDLPGV/VHSLVLTFLVLLLYSSLFPLT/FALGFTITATCGAVGTY
EYSSMAKAKMHYPLSTFSAISGLFVFLALSFGLIRWGHSLGDFPDALPMTLVVWWSITF
RVKSTIGALGLQSTFGLISLVYGIPIRLFHLVLYSGFTHTQEPYGLIWAAGSLATTGKA
DIPGYFGKAGPKNKIIAPQISPNKTVGVFVACGLGATLISGFTFFLQITPRFASYFMPMAI
LPLGLLGLGITGFFGDIIESIFKRDAHLKNSNKLKAVOGMLDTLDSLLSTPIAYLFLLI
TSKEF IG

CPN_0568 657805 658464
 EPPPFSTFATGQPLESFFNGHLTSELTTQEVANAASELSQLPEVRAFMQDLQRRYAQL
 GNCVFEGRDMGSKVFPNADLKITLTSSPEVRAQRKLDLPEGTLSPEQLQAEIVKRRDAAD
 AORAHDPVLPVINGIIVSSDLTIROVLEKILALLFNEL

CpN_0569 658398 650999
 CPG-Glycerol-3-P Acyltransferase
 LFGFDNKTSSGFENFTSKIAMIRICKFTTWAFSLFYKLKVYGVKKNFIKGPAITAV
 NHNSFLDPIALHMCVHECIYHLARASLFNIPWLKQWGCFFVRQDEGNSAAFKIASRLFN
 KRKKLVYIYPEGARSQDQPGKVGIGMMAAKSRVPIPVYIRGTFEAFNRHOKIPIHWVK
 TITCVFGTPIYFPDDIIINPGLKINKETQYQIITNCTMNKIAELKAWYESGCGDVP

Cpn_0570 659044 660789
argS-Arginyl tRNA Transferase
TKLPSKHKHGNRAKGETSPKLMSTLLSILSVCSQAIKAAPFNLEDWAPETPTSTKEHFG
HYGCNDAMKLRVAVLKKAPRAAEAVIAELPQEPFLSLIETAGAGINTFTSPVFNQQLQH
FKYCKALFGQVSPQPKKIIIDFSSPNIAKDMHVVSTIIGSLARIIFSXVGHVDVLRNHL
IGMAITAGFLMILTYLQENPCDVSDELDLTSLYKKAYVCFPTNDEEFKKRSQNVVLAQAD
PQAJAIWEKICTSEKAFQKIYDILDIIVKRGKSYFNPFLPEITDELKKGLLTYSNDA
KCVFHEAFSPFPMVQKSDGGYNYATTDLAAMRYRIEEDHADKIIIVTDLQGSLSHFQLEA
LIAAGLYLQPGIFSXVHGNGVLDPQSKGLTRSGENVKLRLLDTTAEKAAELRHRPE
LTDEAIQERAPFALNIATKYVSDLSHRTSDVTFSEFKMLRFGENTAMFLLYAVYRIQGIK
RRLGISQLSLBGPPEIQPEAEELLATLLRPFPEALESTIKELCPHFLTVDLYNLTHKFNG
FRDHSIHGSPYKSLFLCYALAEQVLTGMHLLGLKTLRIE

CpN_0571 662179 660749
murK-UDP-N-Acetylglucosamine Transferase
murK-UDP-N-Acetylglucosamine Transferase
RNVPDIGVSVLFTLCKSLGAWQWQKTEVLEIYETFEIQCVRPPTFSNVNRIPIILLG
ALLRGCPGVYNTLGAIGERTLHIFMFLGKLOAEQIYSSDSSCYAKAPRGLKGNVYI
LPPYSGVATNLILAAITHAKRGTIVTKNVALEKAILDLVLFLQKAGADITDNDRTIDIFG
TGCLGSVDHTILPKDIEAASFSGVFTSGWRQVFNRNAQKELLIPFLKMLRSITCGGFLFVE
SIEIFFQRPLVGVLVDTLVDPHGGVLVDQGVPAVLLSQAGSSVITHTENVNRLHYLHG
LQHMGAECQLFHQCLSTKACRYAIGNPFSHVAIDGATPLWASHLVTDLRGFAVYMAAL
IEAFGSSILENTLHLDRGYTNWGLKRLSISGAKIOIFDMQREIETSPKSLALRDAISL

CPN_0572 662349 664616
CT556 hypothetical protein
IMAAPIINQPSFTTQTITGCTGTTTTTGVSLGEHSVTTTGSAAQAQTSQTVTLLIADHEMQE
IASQDSGSAVSFEAEHSFTLPLPTTGSVGATQAQAQAGFLPSGRTORDESISSSDGS
S1SRTSNSASSETGSRASPSLQDGLDLSGSEARAGAEFGSGPGPEST1PHYDPTDK
ASLNLFLKNPAVQQMKGTHFVVDARSST1FVRNGDSTAES1KVSNAKNETIK
PADLMECIACFKCVGYETIHSDDTVGRVKPTEEERSGATGNYNHLMSMKFATVYGPWNA
KESSSGYTPSAWRGAKVGTGPIWDDVGGGLGKINMKTPADPDSFINETPGGAHSTST
PGTGPVGATVGVNWRNKLKIGKVDLGGLLGGLITNTVTEEGGGNT1ISTKST2DDKVS
TITSGSQT2IETDIQFPDDPGQDQVNA1PGTNP1PPGPGP1SSSLRLL1TASNALNQ
LQNVQRHLNATVDSNGNSVDLNDQLDGLGNNSGVNFPTVL1PKT2GTDTPDSSQATG
VTEEGGHIRN1IQRNTQSG2SEGAT1PPTPT1AK1VTL3SRGVN5SSSVLPQOVATTI
TPQARTASTT1SIGTGETSTST2TGTCGTG2LKGAT2PTTTT1R2TGT3SATTTT5
PASTOTPOALP3SGTRHATV1SLVRNAGRS1VLVQGGRSQSFP1PPSGTGTGNMGAQLWA
ASQAVST1GOVNNAAATAGSPSSRSSTSPRRK

CPN_0573 665413 664691
VecB family
VEDMAGHSKWANTKKRKERADHKKGKIFSRRIKELISAVKLGGAADPKSNARLRMVIQKAK
KNNINPENIERNLKKAATSAEQKNFEVTSILEYHGCGVGIIVEAMTDKNKRTASDMRIATN
KRGGSIVPEGVSVLNFAKGATCAAKSSIDEEVFSYAIAGAEDLDTDEENFLVICAP
SELASVKEKLISQATGCSEDRLIYLPLRLVDCDEKDGANLALIDWLEQIEDVDVYHNM

CPn_0574 665978 665394
 No robust homology present in GenBank/EMBL as of 11/7/98
 SAERGFRHPIVMGVETVLHNFQRYLSLYKRVFPCRKKTFLSSHRVLRAFSFVVDYCPG
 IYDLQEIYEELNAQLFQGALRLQIGWFGKATKRGKSSVLGLFHENEQLIRHSRLDRQ
 IPIPRFMEYLVYHEMVHSVVPREYSLSGRSIFHGKKFKFEYEQRFPLYDRAVAWEKANAYL
 LRGYKARVCGGYGARA

CPn_0575 ccc524 665982
Phy-Amino Group Acetyl Transferase
LIFVFRWRSFMTAKQNTQILGLEIRYTLPSDATYMLKWLNDPKILRGFPQTQEAETRET
VFWGFGFYRYHSSLTAVYNGNVAGVATLVLPNYPVKVSHHALISIVGEEFRNGKIGTALL
NMLIHLAKTRFKLEVLYLEVYEGNPALHLYQRGFVGEVGRQNRFYKDEIGYLAKTTMEKD

CPn_0576 667543 666494
orfB-Peptide Chain Release Factor 2 (natural UGA frame-shift
MOENLDKRLEALRTEISLAARSL

CPn_0576.1 cc7548
orfB-(natural UGA frame-shift)
MOENLDRKRLRLRTEISLAARCL

CP00577 ~~62586~~ 62486 65815
 WIB (YM74) complex protein
 CCFMSQKNKNSAFMHVNVNLTDLAVVCGKIMPRTEIVPKWVWYIKKKIKQQQKNKRNI
 LDANLAKVEGLSDPDMFQMTFKLSKIIIVK

CPN_0578 **66838** 66838 66835
 1-phenylphosphonolase
 TNSGTVLISRLATLPILAF WAGH IETFWLPITTAIPWLPKKHAILGLRQAQISDII
 HIRVPEKFLKNSKISKINPDLTFVE GDIICARLEDEKERLETTNPTEADIAVFAII
 GHNYDSYISIKNTKKEITPCIEFFKIPFQALAVMXLSSPSRYDNPITQIPMDI

24
LFLYGDHNLGFCAPYLFFFLVLFASGGSGFNQLLVDFWLBESLSFYTHRFDFCKSYPDME
NMEIQAKRKRVEFNLTGEPFKLETLNLYGSGFGLRAKRGVYPVLYALNFGSSCKMDM
DFRGKWNRGSTLTISNQKESINLKLPKDVGVLVNTKTLKGNVCPGSGTFIKQGWGVNKKI
YHNDLWGVSEVTLILNVSSEGGITFS

CPn_0591 680364 681020
yagE family
SLIMRCTAYCTASAINLHVLFLHLKPRYPITLSREYVLANLDSTQASNQLAIFFPFGVAV
TAKMILNTHIKLALTAAGAFATGALHLLHLLHRRPNTASHTNTULAI
NSGLAIVDHTPTPLTPETLPLISLQSMLELMMLANAMLIPLDLKAAVNLE
SGILDPEPDFWDHPETQAIRYDVLSCLDIARINVLIIV

CPn_0592 681132 681461
yidD family
LYSKMFSMKFKRLQQIPVRICLLIIYLYQWLISPLLGSCCRFFPSCSHYAEQALKSHGF
LMGWLISKRIKGCSPWHPGIDMVPKTAQLQVLEPYQEIDGGSSHFSE

CPn_0593 682494 681391
 CT474 hypothetical protein
 VLGA^KMA^FK^RK^RTR^LW^LQ^VL^LIS^LFK^DRY^MGR^PK^IAL^SVA^IIA^KHL^FSG^LPIA^KSS^RK
 V^LSE^DFL^NSE^IQAS^LDL^LIS^LFK^DRY^MGR^PK^IAL^SVA^IIA^KHL^FSG^LPIA^KSS^RK
 TEL^KSG^NSV^RW^LLL^PND^LK^DFP^VLD^LY^LCR^KHY^PYS^KGL^FLA^ITS^KHM^VQBG^NW^DED^CLY^HF
 CS^TPE^FLY^LTR^LRL^TLV^DQSS^VAS^LAR^MV^IRC^GSE^RFI^KPN^EES^TSM^ISA^TQR^KVL
 KSY^LDC^EES^LAA^LLL^LVH^DSV^LHE^FCD^LEX^LVR^LLM^PQES^PY^SQNF^SRLQ^HSS^PRE
 LAC^MST^RVEA^PRVQ^EDQ^EEE^VVQ^DGS^LWL^IIA^KRF^GIP^MDK^IIQ^KGN^LHN^HLF^PKG^V
 LK^LPA^RQS

CPn_0594 682517 684958
 pheT-phenylalanyl tRNA Synthetase Beta
 NTCHYTVQVIVKSLVKTSRLRSMRIPITILLQTYTFSEPLSTKEILEACDHIGIEAEIENKT
 LYSFSGQVATIKILHTIPHPNADKLVRATLTGXEKHQVGVAPNCEAGLIVALALPGAKT
 LFDESEQVITAKKSLGRGVSCQMGCGADELGLDELQIQERALLLEPATPKPGEDLATVLTG
 NLTSELTSLPNLHGASQFLGLAREICHVTQANILVPEKFSFENLTPTTALLMGNDPDICPF
 FSYVVITGISQASPSPKIQSESLQALKQPKINAVIDTITNIMLSQLQPHAYDASHVALDS
 RLVKELTSPESLTLLNGETVLPGSVGVVRRDDSLQLGGVGMKAPSPQETTTTIVIKA
 AYLFLPEALRASQKLLPIPSESAVRETRGIDPONVVALQQAIIYLIEFPEATISPIYSS
 GEICRELKALREKPTLQRIILGKSPSIEILLSQLKSLGFSFTPOETSLQVVLVPSYRHDI
 EEDLDVEEIKRTSESNITQNPVSCYPTPIYKLKRETAGFLANAGLETFDPLDLPTETVA
 LTRKEKEEISLQGSKHTITVRRSLSLPGLLKSAATNLNRQAPSVQAFIEIGTVYAKHGEQYQ
 ETQTALILHTDEGSRSLWPKPSSLFSYSLKGWERLLYHHHSLDCLMTLLESSACEFHYQ
 QGVGLRIHKQSFATLQGVHPELAKQAQIKHPVF AEELNLDLLCKMLKTKTKLYKPAIYV
 SYFRDLTLTVEPDI PANLLRQVALLHSGKSWLSEVSTIISYQDKSLSTNNKNSLRLVFQD
 YERTLSNODIEEYCRVALNELLTDTKGTINS

CPn_0595 684943 685926
CT476 hypothetical protein
RDYQFMKQLGFCVCFVMSACSAAYSPRRQDPSVMKETFRRNNYGIIVSQGEWVKRGSDGTI
TLKVLKNGALTHEYVSGLLHGEITLTPTHTALDDVQIYDQGRLVSRKTFVGNLPSQEE
LFNEDGTFVLTRWPDNNDSDITKPYFIETTYGHHVEIGSYTFNGKYSSSIHNGEGBVS
VFSNNNLIILEETFGNVMGVKTYTPIYFNRPDSITHYQNGPHGLRLTYLQGGIPINTIEE
WRYGPDQDGTIFVFNKGCTSEIAVFKVKGEGLELRVNEQEIVAAEEVSWNRNLDHGKERKI
AGGIQKHWEYTYRGSRVSKAKENLAAG

CPn_0596 685930 686457
ada-methyltransferase
FAVMADDTLPIKLMNSLSQACSEGLLIAKYPPLQIVVHFDNNLVVKTHTLSVAPVFSCLF
CPAARIAKMGQIIVLWCSRYANKEHHPFCHSVFAKDLIPSYQYLEILNCVAIEIPFGEQQTYAE
IAKKTDTHTPTVAGACRONPFLLEHVPFCHRVVSHGERNYVLGVPVIEILLKFENSY

Cpn_0597 688215 696479
oppC-Oligopeptide Permease
RMYKFFSPYQRFLLSAYYKNLLASLSWKFIFVSVALIGYALFLFASSKPLLVTHGHEIFFFLL
LVLPFGGYTQKPDVLFNNVWVTPFFPILSGFKLTGRWLRLLWLLGLCISQCMIFAWAYSQ
KVQPDALAEMLKMKRAEKRSINVSVMSEMLPKDTRWEMERYMSYTEQGLLIQA
KRYKQKQASVKYKQVAFEKRSQSPMTLRHLEMNQEGICLRKQORVDMKQRPYMAQQA
WNRRATDNYRPFMLALTRHEHLRLANVNNMGQPDLCIAYANVAKKRAEPYKSKLLEIRF
LEDYAKRLASISQVQDKRLWTEKESEDRLINPFFSGFWEDDAGSEMNKYPPWQOL
SRPVRKDLLALVGLFRIGLAVVAGIGITIALAGISGLVSGYGGTVDMLSRTEIWE
TMPLVFTLLMLVISITQOKSLLLNTVLGCGTAFGRIMRVYRLEVLKQRDRGYVAATNLGY
SHYYIMVHQTLPNATVPPVISLVPFAMAMISCEAGLTFLGLGEESSASWGNLMRGVGTGF
PAESAVLWPPAILLTMLTIALITGDCVDRDLPRLDS

CPN_0598 689712 689219
OppB-Oligopeptide Permease
EEGSGVSLYIKLRVLVLIPTLFAIVSIN FVILNAA PGDVLEEKS RDALGEAGKSGDKMRSY
GKDPDYLQRFREHYGLTLPIFPNTRPKITHKKTQALQELANANNTTSPAKNAASLVYWG
QCAKPYLVLRDAGEADSRDDGKTYRHIAADLPIRGGVGLQFVGNLSPEQRAQNKIEAEN
FVLRLQNEDDLTKVEALKGWDPDHGTEVFCYSQKQWKTFFLETRFARYMSVRLNRD
GFTLRNDAAKHTYISEVYIKRLKSLVLSLPMTPVGVLCQIFGMIALKNRNWIHSLRFL
LFLILFIPVFAVAPVILNDLVPINKTIPPTQIPMPYSGLRSPPEVFNELSTGRIFDLVSH
FGLPFCFAVSAGALAAQSRLSRSIFLEVLQDQICAAKAGRLRVLIDLYKHVGKNAASIV
TSLAASLGLTLLGALVYVETLFNIDGPNGFYQAILNRDHNVVLFSVLVGSALSVLGYLLG
ICVYLLDPTLGGECRRGI

CPN_0599 691823 699682
 alpha-oligopeptide Binding Lipoprotein
 KRCEGHMYKRCVLKILKQIVAGSLILLYWSSDLLERDKSIKGNVRDIQEDIREISRV
 YKQKQTQQAIPAAKVMMLAPKLVRDEAFALALPGDPSYNLLSDLYKQKQTEPELLGTNFK
 LKFLLRTHATFPHNLFPGNFGVVGFGYDLICPLGASHPDVKYEEFSDPLAVKITEEHLV
 EDCGCGKEFHILYLRNFVWRPILPKALPFLVWQVDEVFVRHPVTAIDIKFFDYDMMVNVYV
 TNPMAVALRQSYEDVYLYSVENCLLKVYPHKAITLYINDEIKERKLYI:AFS:NTSLQGL
 TRFYQYQITANGKIIIDENIDI(PFNG)YIAQNFNTMHWANNLYV:YAYYFA:MDDKEIVF
 SKNIIDVYDPLALIDKRFYVYFTE:TLIDQETKIKIDVLPFNORDINYYFMKSAAYN
 QKAPVYAVRQTVSADRAYTYISANKFELPQVQVVRVAMMMAIDRERILDQGLDQKQYT
 KTFEAGSDPSYNNKLEWYIYDPAFAALLLIYALIDYIKDIREKVIDVLYVPRFRLE
 YVYQYTAHTIDAYVYACKETIGIDVLLIYIMADLQAFDEPNFADALLMKTALIPED
 RALMIGDGLAMKFA:ANVGFYHIEFAEPIIDRI:YEEYDKEPHRIYHRFPIIDREAFYA
 IL:SPHCLSLYKIVKNIYVTHPIYTHIYD:KLVYVNVVMVYKKEDDIL:LTG

Trn 9600 69156 691827 691827
do robust homolog present in Genbank/EMBL as of 11/7/98

FELWEE:QENEIVEQKKFQLPPPAKLIGEVIQTQVDPVVTSSADLNESLQALVRESSDL
INALLJADDAIHFPETEEPTSAFEESSAMFFPETSSATEEE

CpN_0624 718018 717011
gapA-Glyceraldehyde-3-P Dehydrogenase
AMKVYVINGFRIGRLVLRQILKRNSSVEVLANDLVPGDALTLYLFKFDSTHGRFPEDVRC
EADHLIVGKRKQIFLSERNVQNLPKDGLGVLIETCTGLFTKKEDAETHIQAGAKRVLLIS
APQKQDPTTFMVMVNHKTENPEKDFITONASCTTNCLAPIAKVLLDNFGITTEGLMTTVHA
VYVPLDTHLEFTHWQVHKAQNLVLAQVAVYVLADEKRLVLYMAFRVDFIDV
VYVLLTTRLEFSTTQDDKAMKQVLTLEKRLVLAQVAVYVLADEKRLVLYMAFRVDFIDV
IALNDRFFKLVAWYDNETGYATRIVDLLEYVEKNSK

CpN_0625 718488 718060
r117-L17 Ribosomal Protein
VMQHARKKFRVGRSSSHNRCLMNLKSLIHYERIETTLPAKELRRHADKMITLAKKNS
LAARRIAIGRLMVRYNKLTSEARQAKGGDTSVYVNDRLVNNKLFDELGNRFVERKGGYT
RILKLNQRIGDNAQKCIIEFLAS

CpN_0626 719670 718495
rpoA-RNA Polymerase Alpha
WLPKAKKAQSVVLGKEKGMSDNAHNLVYDKFLEPAVKMLPVEGLPIDKHARFIAEPLER
QMGGHTLQNALRRALLIGLEAPAIISFAMTGVLEYHMAIEGVIEDVTNIILNLKALLKKY
PMQDSSLGRITTOVLKASISIDASDLAANGQKEVTLQDLLQEGDFEAVNPQVIFTVTQIP
IQLEVLVRIAFGRGYTPSERIVLEDKGVYIIVLDAFSPVTLVNYFVEDTRVGQDITDFDR
LVLIIVETDGRVTPKEALAFSTQILTKHFSIFENMDEKKIVFEAISIEKENKDDILHLKI
LGNETELSVRSTNCLSNANIEITIGELVIMPEPRLQFRNFGKSLCEIKNKLKEMKLEL
GMDLTQFGVGLDNVKEKMKWYAEKIRAKNTKG

CpN_0627 720059 719640
rs11-S11 Ribosomal Protein
FLIRSRVLVKNQAAKKSVMKQKKNIPSGVVHVKATFNNTIVSITDPAGNVISWASAGK
VGYSGSRKSSAFAATVAQAQAAKTAMNSGLKEVEVCLKGTGAGRESAVRALISAGLVVSV
IRDETVPVHNGCRPRKRRRV

CpN_0628 720461 720063
r13-S13 Ribosomal Protein
DAYTILREQAQMPRIIGIDIPAKKLLKISLTYYIGIGSARSDEIILKLDPPEARASELT
EEEVGRNLSLQSEYTVEGDLRRRVQSDIKRLIAIHSYRQQRHLRLSPVRGQRTKTNST
RKGRKRTVAGKKK

CpN_0629 721881 720487
secY-Translocase
KIELFRPYMTTLRQFFLITELRQKLFYTFALLTACRVGVFIPVPGINGELAVAYFKOLLG
SGQNLQOLADIFSGGAFAQMTVIALGVVPIYASIIIVQLFLVFMALQREMRESSDQGRK
RIBBLTRIFLVALAVIQSLFLAKFALRMNLTIPGIVLPTLLSKLFGVPWFYITTVVVM
TTCTLLMMWIGQISDKIGNGISLIIALGILSSFPVSLGSIYNKLNLSQDSSDLGLIS
ISEHALVVFVFLITILLIEGVKRPVQVYARRVIGRREVPGGSYLPKLVMYAGVIFVIF
ASGLLMFPATIGQFIASESSWMKRIAAALAPGSLVYSICYVLLIIFFTYFWTATQPHPEQ
IASEMKNNAFIPGIRQKPTQHLEYTMNRVTLGALFLAATIALPSLLGCLLRVDSNV
SYELGGTAMLIVGVVLDTMKQVDAFLMRRYDSVLKTDRTKGRH

CpN_0630 722316 721885
r115-L15 Ribosomal Protein
MKIESLPDISERKRRKLLRGPGSSGHGKTSGRGHKGDGSRSGYKRRFGYEGGVPLYR
RVPTRGFSHKRFDKCVETITGRLAEFQEGEAITLDALKAKKAIARQAVRVKVLKGLD
EKTFVQQTAVVLSQGVQNLGKIT

CpN_0631 722812 722312
rs5-S5 Ribosomal Protein
EBNLSLKNSHKEDQLEKVLVNNRCSKVVKGGKRFSSALILVGDGKGRGYPFAKANEL
TDAIRKKGAAKNLMKIELEDGSIPEHVLVHHDGAQLLLKPAKPGTGIVAGSRIRLLI
EMAGIKDIVAKSFGSNPNMQVAAFKALTGLSPRKDLLRRGAIND

CpN_0632 723354 722827
r118-L18 Ribosomal Protein
KGLISSWLVNLLQVFAPNVLLNLKIVREFVMKMMNSVVKLVLRQAQKNRSVMESSLCK
KSLIKRRRALRVKRLKGSPTKPRLSVFMKNHIIYVQLIDDSIGKTLASVSTLSKLNKSQ
GLTRKNQVAVKVLGTQIAELGNLQDLRVDFDRGPFYHGIIVSMVADGAREGGLOF

CpN_0633 723760 723209
r16-L6 Ribosomal Protein
SMRKAKEPILLPGQVEVSIQDDKIIVKPGKSLTQKSVEKEITLKDINSIFVHAAPHV
DRPSCMQGLYVALISNMVQGVHLGFEKRLKEMIGVGFASVQGAFLDLSIGVSHPTKIP
STLQVSVEKNTLISVGLDKQLVGEFAASIRAKRPPEPYKGGKIRYENEYVRRKAGKAAK
TQKK

CpN_0634 724215 723787
rs8-S8 Ribosomal Protein
ESSIKRKRIYMGMTSDSADLTRLIRNALMAEHLVVDVEHSMKREIAVKILKHKGFAVH
LVKEENRRAMRVFLQYSDDRKPIHQLKRVSKPSRRVYVSAKIPYVFGNMGISVLSTS
QVMGGLSARSKNIGGELLCLW

CpN_0635 724763 724206
r15-L5 Ribosomal Protein
GERKANMSRLKFFYTEEIRKSLFEKFGYANKMQIPVLKKIVLSMGLAEAAKDKNLFQHL
EELTMIQQQKPLVTKARNSIAGFKLRQGGIGAKVTLRGIRMYDFMDFCNIVSPRIRDF
RGFGNKGDRCCYSVGLDDQQLFPEINLDRVKRTQGLNITVWTTAQTDDECTTLELMCL
RFFKKAQ

CpN_0636 725100 724750
r124-L24 Ribosomal Protein
PKEKEVMKKONIRVQDKVFLACNDKCKEKGVLSTEDKVVVVEVNVRIKNIKRSQONPK
PFRIGIEAPIHISNVRLTIAGEPAKLJVKVTQGRELWQRRFLGTSQLYRLVRGKKG

CpN_0637 725473 725099
r114-L14 Ribosomal Protein
IRHIMIQKQQLKQVADNTGAKKVKIKVQLGQRRRYATVGVIVGVVRVDFVNNISIKKGDV
IPAVIVRTKRRIITKKIKSTIKFDINSCVITDOKGNFKQTPITGVGAREIDRGFTIKISL
AFREVI

CpN_0638 725771 725490

rs17-S17 Ribosomal Protein
NKKEKVKMSAGEPPGSRKVKIGVVSAKMEKTVVVRVERIFSHPPYLVKVRSGKKYYAHT
ELKVGEGDKVKIQETRPLSKLRWRVIEHVGVS

CpN_0639 725979 725743
r129-L29 Ribosomal Protein
ASGGKINMAAKDLTLTQLRKSDDDLDAYVHENKALFALRAENLLQNKVVKVHMFSTHK
KNIRALTLVQEPFKVHG

r110-L10 Ribosomal Protein
IIMLMPKRTFRKQKQFAGLSKQATFVDFGEVAMOTLERGWVTSRQIEACRVAINRYL
KRRGKVVIRIIPDKSVTKPAETRMKGKQAPDHVAVVRPGRILFEVANVSKEDAQDAL
RRAAALGKIKTRFKRVVERV

CpN_0641 727092 726409
rs3-S3 Ribosomal Protein
KGRRIIMQKQKCPIGFRTGVTKRWSLWYGNKQFEGKFLIEDVIRIQFLRKKPSQCGAAGF
VVRNMSGKIEVTQTARPLVIGKGAEDVLLKEELRALTGKEVWLEIAEIKRPELNAKL
VDNTRIARQIERRVSRAMKAMQSMADAGAVGVKIQVSGRLAGAEIARSEWYKNGRVPL
HTLRADIDYATACAEITYIGIKVWINLGENSSSTTPNNPAAPSAAA

CpN_0642 727440 727096
r122-L22 Ribosomal Protein
RRHSMFKATARYIRVQPRKARLAAGLMRNLSVQEAEEQLGFSQLKAGRCLKKVLNSAVAN
AELHENIKRENLSVTEVRVDAGPVYKRSKSKSGGRSPILKRTSHLTIVIGEKER

CpN_0643 727725 727450
rs19-S19 Ribosomal Protein
ETRIMGRSLRKGPFVDHLLKKVRAMNIEEKKTPIKTWSRRSMITPEMIGHTFEVHNGKK
FLTIVFSETMVGHKLGEFSPTRIFKSHPVKKG

CpN_0644 728594 727722
r12-L2 Ribosomal Protein
FIREINSMFKKFKPVTGTRQLVLPADFELTTRGELRGTGKSRSLRPNKKLSFFKXSSGG
RDLGHISCRHROGGAKQLYRVDFKRNKDGITAKVWVVEYDPNRSAYIALSYEGEKX
YILAPKGIQGRDVVVSSEGGSPFKGCCMTLKSITPLGLSVHNIEMRPSGGKLVRSAGLAA
QVIAKSVYVTLKMPSGEFRLNEGCRATIGEVSNADHNLRVLDKAGRRRMGVVRPTVRG
TAMNPFVDHPHGGGEGRHNGYIPRTPWGKVTGKLKTRDKNKNKNIWIKDRRK

CpN_0645 728933 728598
r123-L23 Ribosomal Protein
DMKDPYDVIRKHYVTEKAMLEHLSAGTGEGKKKGSFCKDPKFVFIIVSHDATKPLIAQAL
EAIYVDKNVKKSVNTINVKQPARMFRGRKGTSGFKKAIVTFYQGHSGV

CpN_0646 729636 728950
r14-L4 Ribosomal Protein
YREDLVLLSKDFSGNKGVEVADSLFADEGDGLQIKDYIVAIRANKRQWSACTNRN
SEVSHSTKFPKQKGTGNARQGLASQFRRGGIVFGPKPKFNQHVIRINKERKAAIRLL
LAQIKQTNKLTVVDDTVFDALTPAKTQSALRFLKDCNVECRSILFDIDLHVEKNENLR
LSLRNLAVKGFVYGININGYDLSAHNIVISKKALQELVERLVSETKD

CpN_0647 730490 729657
r13-L3 Ribosomal Protein
YLEYFSYCKNLPPLITCFPIFLRENFLFLENSISKILSRFVSLFLQEEKSKLLMDKFM
RSHISVMGKKEGMIHIFDKDGLSVACSVIRVEPNVVTQIKTKESDGYFSLQGAEMNAP
AHTITKRVSKPLGHLRKAQGRVFRFLKEVRGSEALNGVSLGDAFLEVEDVSSVDVR
GISKKGFGQGVMMKFGFRGGPGSHGSGFHRHAGSIGMRSTPGRCFPGSKRPSHMAENVT
VKNLEVIKVDLEKKVLLVKGAIPEGARSGIVIVKHSRT

CpN_0648 731636 730605
CT529 hypothetical protein
FFFKPKCEVKMATNARSAGSAASKMLLPVAKPEAAVSSFAQKGIYCIQFFFTNPGNKL
AKFVGTSLDKDCFKLSKAVSDCVVGSLEEAGCTGDALTSAARNAQGMKLTTRVEVALANV
LNGAVPSIVNSTRQCYQYTRQAFELGSKTKERTKPEYSKMLTRGDYLLAASREACTAV
GATTYSATGFGVLRPLMLINKLTAKPFLDKATVGNFGTAVAGIMTINHMAGVAGVAGGIAL
BQKFLKRAKESLYNERCALENQSQSLSGDVLISAERALARKEHVATLKRNVLTLEKALEL
VVDGKVLIPLPITVACSAAISGALTAASAGIYSIWQKTKSGK

CpN_0649 732672 731710
fmt-Methionyl tRNA Formyltransferase
NLNKVYVFGTPTFAATVLQDLLHHKIQITAVTVRVKDPKRSQALIPSPVKTIALTHGLP
LLQPSKASDPQIEELRAFNAVDFIVVAYGAILRLQIVLDIPRYGCYNLHAGLLPAYRGAA
PIQRCIMEGATESGNTVIRMDAGMDTGDMANITRVPIGPDMTSGELADALASQGAELVLI
TLQIESQQLQVLVSDAALATIAPKLSKEEGQVPMWDPKAEEYAHIRGVTPAPGAWTIFS
FSEKAPKRLMIRKASLLAEAGRYAGPTVVVTDROQLAIACSEGAICLHEVQVEGKGSTN
SKSFLNGYPAKKLKIVFTLNN

CpN_0650 733513 732665
lpX-Acyl-Carrier UDP-GlcNAc O-Acyltransferase
SRNMAISHTAIIEPGAIGKDVVIEPYVVIKATVTLCDNVVKSAYIDGNTTIGKGT
TIWPSAMIGNKPDQDKYQGEKTYVTIGENCEIREFAIITSTFEGTTSIGNNCLIMPWA
HVAHNTIGNNVLSNHAQLAGHVQGVGYAIILOGMVGHVQFVRIGAHAMVAGALSIGIRRV
PPYITGSGNPYQLAGINKVQLRRQVPFATRLALIKAFKKIYRADGCFESLEETLEEYD
DIPVEKNFIEFCQSPSKRGIERSIDKQALEESADKEG/LIES

CpN_0651 733975 733517
tab2-Myristoyl-Acyl Carrier Dehydratase
MNQPCVILKRELLDLLPHRYPFLVDKLSYDIEARSITAKNVTINEPFFMGHFPNAPI
MPGVLLLEALAAQAVLIGLVLENDNRKRIALFLGIQKAFRQAVRPGDVLTLQADFSLI
SGPQKAWAQAQVPLVTEAELSFALVDKESI

CpN_0652 734880 733990
lpX-Myristoyl-GlcNAc Deacetylase
KRNITLVDLLEAFMYLERPOTRLKREVRVQGVHLLQKSTLHLOPAQNTNYIVTQRCQ
AGNINYNPALLPHVYTYRITTLGRQCAVATVEHLMMAALRSNNIDNLTQCCIEEIP
GDTQNVVVELLDQAGICEQEDKVSITRLTPVYVYQHGDIFLAAPFDELKIVTYLHYPO
STTGTQYKGLVINEECPROETAFRTTALYNECLMEKGLIOGQGLDNVAVFKDQXII
GRQLRFADEPVRHKLLDGLDGLVGRPFVAVHVAQGGHSSNIAFGKKILEALFL

CpN_0653 736499 734868

103

CT668 hypothetical protein
AFKTVKRFECFMIDPVECFPHNLDGDAEAQSIITQNGSTFLA:SELKKDITPFAIGSYAAPKD
TLTVQGVKFNPMAMQDQGNILIDPELQEALEGEELGEQIINLKGRLWDFRSTFEDSQT
AQFADEDFQAVGVIIIDLINDELNTIAEHTQDARDFKDEEGSVSTRIIDWGSSEEVNLR
ALLAFESDRDGNRESLANGLQVYAVQVQRAELFASIVGTSVSSVKTIMTTQLG

CPn_0703 796203 795742
CT667 hypothetical protein
CT667_0703
OVIANKIESTRALQSYLDWHDITLVAKJACPL

CPn_0710 796482 796210
CT666 hypothetical protein
RSRGEKSMATNKSCTAFDFNKMLDGVCTVYVKGVQQYLTELETSTQGTVDLGTMFNLQFRM
QILSOYMESVSNNLTAVTNEMITMARAVKGS

CPn_0711 796791 796486
CT665 hypothetical protein
TTINNQVLGFINYLYLGRYSMFNMENTAKEEKN SQPLLDLEQDMQDHRAQELKASVQDK
VHKIHALIREGSDKESFGQOOSLLAGYVALQKVLGRINRKMI

CPn_0712 799315 796781
FHA domain; homology to adenylate cyclase)
MAVRLIVDEGPLSGVTFVLEDIGSWIGRSSANDIPIDPKLGASQAIINKTDGSSYYIT
NLDTIPVAVGVATQFTDLQKNDTILGSSNOYSELSDPEEDPDLPAYDIEDIPFEENFND

SGDLSDSNEQKDLPEPRTSETNHSPPKPEKRLTKQDSSSPVITSGDQFLADAF LASAKAE
QKQNPRAKVAGKLGKSSNESLNPKQENAKDPSKGEERNTKPNQNA TMDGNAS PRQDPQPK
SAEPLSKNTARTDET LKPNKEVPEEKANKKATP DSPKPKQDPEEGSKBEGSKT EATPLDSQ
KESDEKAEAEAFVQEEENITDENKDSASAANDADTASDHTADENKTKPKVKNESKA
VLSPHQLTGDLFRFDQT IPFAEIDDIKAKNSVGLTQPSRFLVLKVGANIGAEFHLDSGK
TYLHGTDPPTICDITVNDLSVSHQAKRITVNGNDGIL EIDLDSKNGVIVEGRKIDKTSITL
SNQVVALGTTLLFLLIDHHPADTIVASLSPPDYSFLGRQDQAEALQEQEAEQEEEKQKRA
TLPAGSEFILLTFVGSGLALFGIGTASLFTHKVEVPLENDIYEDLERQINQFPTVRYTFN
KTNQSLFLIGHVKNSTDKSELLKYVDLSFVKSVDNVDIDEAVWQEMNLTLSKRPFPKG
ISMSPSPPGFKITGYTVQTEBAQACLVDFYJNLHFNYSLLLENKVVYQTMKLAAIGHLLQ
GGFANIHAVFNVEJGVLTVGVNNDCEAFRAVQGLSGTIPGVLVKNFVALPLPAEGIID
LNLRYPNRYRVTSYSGRYSEINNVNNGRLITRGDVIDGMTVTSIQNAPLELEKGLKYIA
TDYNK

CPrn_0713 799817 799332
 CP663 hypothetical protein
 LDLLKEKAGFRNEIVSIPOGKTITIAALENTSMLEKLIKNFATYMGITSTLELDADGAYV
 LPISEVVKYRAQGDYDNEIVLSASLGLAPPASDTAKLYLQMMIGNLFGRETGGSGALDLS
 EGNVVMVRFSGDITTTDFVRHSEFMNFSETLSLDLGLGKO

CPn_0714 801125 800091
 tRNA-Glutamyl tRNA Reductase
 LKMLGCVGISIRLAKTGERRAQYLGSEFKNLPAQRFGLKGGAFTPLLTCHRA
 ELLYYSESPIEAQAALLSEATSGRQPRYRHGKCTFHLFGQTSGIDSLFGFETIEQQCV
 KTRFLGKSGKERLEPFDLFDLFGQALKKEGKEVRSRIGFPDHPQVITSEVUQEILLSDYKSI
 NNLYFVGYSIDINKVAAYLVQHGHYRITFCSRSQAPVYRTLSRTELSPROPYDVIFGFS
 SESASQKDLLSDCSLSEAIKPRIVFIVNPVPTFLWKECTPFGVYLDIDFISECVQKRLOCT
 KQKYNKADLLSSCAAKKQVEYFKKSSITGROIISPRISVGLY

CPFN_0715 801636 803462
gyrB-DNA Gyrase Subunit B
KFNKISHHMAAYTEASILSLASLDHRLRAGMYIGRLNGNSQKEDGIYTLFKEVVDNNGIDE
PIMHGHSKLSKISASDQKSIISQDQGRIGPLKIDCVSKINTGAKYATQVDFHVSVLNGVG
LKNALVNSEISVSFVSRRKKYKYLTHFRHVLQSGKSGSTGPKDQDPTQSTPDPSP1PEFT
FNHFDLKKDKIRQYTYLHSGLEIRFNDFEITFASHNGLKDLFDAEITEPPLYSPLPFQNNED
IFSHLEGNTRYEVSFVNGQETLDGHTLTAHKAIVGVNVEFFGKTPFVSNDIREGIVLT
KAIATIASPIFESQTKNKLNGTIOIRSSLLKDVKEAVLQALRKDXVPELLEKIKFNKTR
KNIOIFQKDLQSKQKKVYKIKPLRKDFKHYNDRSLYGEASDIPVTEGSEASALSRN
PLQIVNVSFLRKGPMNVFSLTEETMYKYNDELFLYALATGIGTONEIHHYRNKVKILATADV
DGMHIRANLITLFFTKLLPLVNNHLLFLETPLFGVRKNNTTLLYYSEQEKQKQOFGK
KDDSLSEITFKGLGEISPKEFAAFIGPEIRLTPVTITSLSEISSILQFVGMKNTKERKQF
TMDNLITDF

CPN_0716 803466 804902
GyrA-DNA Gyrase Subunit A
FWRIYVALFRTHFMHVAIVTIGLEPIHILDLKVPQRRRLGLDVLMDGGKGVAAVNIAG
RTATMAVHEHGDADIVLEALVVLVANGVYLDITDGNQFNPLGDPHAAARYEARLSPRLARET
FNTDILTAFHDSYDQGREKPDILPAKLPLVLLHVGQIAGVMTKIFPHNFAELKQAIAT
LNDKKFTFVPSDFSGGKMDPSEYDQGLGSLTIRASIDINDKTLVQKICPOSTTETLIR
GYNKAEARGGTIKITIDQFSTOVGHIEIKLPGSGRAKEMLPFEHTEGVQVILYSKPTVI
IENKPVKCSISEILKLHVLQYGEKLEELLLQEQILTLDHYHKLETYFIIKKHLYDSVRE
LLEIKNISADDDQLVAVLHPLHRELATFVPTKDTOSALSTIKKLCPNEBAECTKL
LAIEKKKAAIOKDIARKEFVTVYKLCGLFHHCHLGFPEKTIIDNFKATKSLIKOOTI.L

CPn_0717 804968 805306
T656 hypothetical protein
IRIKFIDITIVVRMEPRHIVIRKPEPKAPDVEKPGVPEYMTMANTPTFEGPVKTLQDL
RRALIEORGAEEGQKMYDNFIOSLILSTFGLVHKMDPAOKASKRMRSVYKEO

CPn_0718 805300 805626
 T657 hypothetical protein
 YAMMGFTYFLALFVDRMLQERFLCSPKRWAPFINSPLYLT/LIADHDTTPYLAKNLDKFPLP
 YQWCKTTLHVGSLLKIFLCSGLSSRLLLAKTKFEILTNDLYAONI

[illegible]

14
 LURFTVFFLPKILQGLLQDAFIPHEEFLRAQSLDRAAFFERFERLRLKSTIITFLTE
 AVWVVVLQVVECTVYMLLTLMLLPQIFLPMYNNVAGLHCENKFTVGLAPVNNI
 WIFFVILAAHSDPREIIGLGLVWIGFFVEVITPGVWVLLLEAKQPEQHDVRLLL
 APLGLSGTISFQNLNLSIDCLARYVHEIGPLVLMGLSKLYQVLPILHFGFQSVVLLPA
 SRCVQREDSHRLGLMKFVLTLMMSVIMITAGLLLLDPCVRVLYEHGLFGPOSAVYA
 VRVLGYGACITFMALAPVLGVLYFAQRQVAVPLFIGTATANNIVLSVLGRWLVKDV
 GSYATSIITAWVLQVFLAYVSSKRLPMYSKLLWEGIRTSIKVMGNTLMACMITGLNLT
 OTTVYIFLNLPLFLAWPLGSIITQAATAFIEGESCIFLAPLFGFAKLLRVEDLINLASFY
 WLLRLTGM

CPn_0731 821494 421760
No robust homolog present in Genebank/EMBL as of 11/7/98
VAIAISRNPIVIRLQK/PDNIILKIERAKETSLSLFLLIKPFPSPPLKQDYLFDISPYTSSE
ITIGGSYFKLNKASLSSSTLRLSISIIIS

CpN_0732 822092 822976
nfo-Endonuclease IV
NMFKVLPPSPISLLGAGTSAGGLKNAIYEGRDIGASTVQIFITANQRQWRRLKEEVIE
DKAALKTEKLSYIMSHAGYLINPGAPDVLVEKSRITGYEITQLDITLGSFNVHPGA
AFKSKCEDKMNDVSYSSFSQASLPDSSPLVLVLTITAGGOTLIGNSFEILGVQLNKLK
QIPIGVCVDYCHIAFAGYDITISPGWEDVLNFEFDYVGLSYLRAFHLNDSMFLPGANKDR
HAPLGEYIGKESKFLLMTDERTRKIPKYLETGGPGPNWKEIGELKFKSKNRDS

CPn_0733 823739 823101
GLYKARYCVGPKNVRARFGINAFGRSRNPCLKKPPHGOHMQRKKKSDYGLQLEEKQK
LKACYGMIMEKQLVKAFKEVIHQGNVAMQLFERFECRLDNMYRMGFAKTIFAAQQLVA
HGHIIVNGRRVDRSRFLFRPGMKQISLKEKSKRLQSVKDALESKDESSLPYSILDTGFK
GELLVSPEDDIEAOLPFINISVCEFLSHRT

CPn_0734 823863 824915
ycaA
QNTKEHFSSNGNLQCNFYQDYVRAEIMEKKYALYYITRVDPNHEEIALHKHFLLED
VDSRIIRYSIQGNIQDQSGYEPHEVIMYGLKLRNFKSIKFKIHIIKENITPRITVKYR
KELAAEQCEVDLSQAQKHISPOEWHEKQJENRCLILDVNRNMYEKGIFGNATPLDQITQ
REFPEYAEKLAQKQDEPTTPVMYMGCTGGICRELYSPVLKKGFKVEYLDGGVIAYGQV
GTGKWLKGLFVYDDRLAIPIDESDDPQVAFIEACCHCQTPSPDAYNCNANTCNALFLCCDE
CTKHQOQCCTFFCSQPRVRYKEDSBNKPEERABHCFISSENSFASCCIL

Cpn_0735 825680 825003
 *Uridine Kinase (Uridine Nucleophosphokinase) (Pyrimidine
 Ribonucleoside Kinase)
 GEFKFLMPLMNIIGITGGSGAGKTLTQNIQFQDVSVICQNDYKDRSHYTPPEERAN
 LWDHPHAAADLNDITDKRLKNNEIVQAPVDFVLGNRSKTEIETIYPSKVLVVEGILV
 FQENQLRDLMRIDFVDFDADERILKMRVRVDEQGDSDVDCIMSRYLMSMKPMHEKFIEP
 TRKYADIIVHGNYVRNVVTNITLSOKIKHLENALESDEDTYYMVNSK

CpM_0736 827731 825992
cyeD_Efflux Protein
RGELKKLARQCLVAFMTVSVKKKSFRAVLVTHFLTIINDNLYKFLLAFFLLEGGKLTENA
KLCVCSFFFPALPLAGSLADGAPFOKRNILTRF IEILCTILGTGYFFGIQSVGGV
YVVLILMACHTTIFPGAKLGIPELMPSEQLSQANGIMTAATYTSILGKCSAPLLVDVT
HRLGVSVVPTLKCMTVTSIISTLSFCIRSPNSNVKVKGTILVSVFKDLKVKGLTRMIH
VLTVSIFLGSFGLKICAYTOLIEIPFVEFKLPYKHGYKAYFLPVALGVGGTSGYSITGKIS
KXDKIKIGYVPLAIGLALFVPMGLYAFACLILPVLFFLLALGFLGGVYQPLHAYVQYASP
EHKRRQGLAANFIPLDFGVLYAAGIVPVLGNSLGLDPSPTFFYIGVFLAVSIITWLTIWR
EHVYRLQLIILRRQLGYLQVLTGHSSPKVYFVAVQSYREIRRILGWALTITVRSRNIILD
KVLPGVRWALLSCVPTVSVSSVRDNEQAQDAWVLQANHLKTSKKKFPDVSVVCGLGP
KNVRFSLTLEOGDILHPQILVOREKGGDRIWTVLVFPHA

CPN_0737 827469 830756
 reC-Xcodeoxyribonuclease V, Gamma
 KRKSAKLPAASGASGRKGRAKKLLQERIFAFSVRVLPNKRKNAKRNLYKLSFIIVRKCVVT
 WAKNQLVFLLETVMNATKHCRASFSPNSPHLLAQLEDTISHQPKFTKRWILVANATGHW
 IWKINQLVFLFETFMGTISITFASDSIVKHLFGSSGSGSNPNIDYLTLLNNILNIEEIS
 KAKFTENGREGLFSPPTYETTKLAAAFQKHTFSQRTPKNASHYQELFOILIESHFSYEE
 MPTTLLNNRKTQEEDCSLHIFGYAHLPKHIAEFTINSLTYFPVYFCFSPREYQDGLLSD
 LADFDWNQGLDPSPIKNAAWEHYLSDRQALLANLHKSSQNFLLDREIDYQEMFLPSK
 HDHSLGYSIQNSLIDLKPTSPQDSFTKQTCIYRALNIPREVQEYKVTLLHGRVSEY
 EFTLLSHSEYKYVHALLFNPHBVIYFTDEVDPAEDLLNNKLLLSLITQOGLDHYLIT
 QLLTHLOQLOPQIDONKVPVYLKILLSSEMGISKKDRASQKMGALGDILEEYFPHQEGG
 RVSOVEVWETVPLTYIYQERINLYLSSGSSYEDLPQNVFSCLEKIFVLSPETSFTIT
 LRNSLFTPTFATSSCSLLFTDPCFLDLHLHFHKPNSLYDKPGPYIGLSLSLILPKGVYI
 LGANKNTTSETIDLLNRTTHLEELFASSTEDENCFHFLQILVSTKHELHISYISAAQFN
 PSPFNNLHKSDILDLVPTLPTQYPLSAFFKNKSTQSEYNYSALHAAYFSKALLPSL
 IPTVQKNVQLHQLSNEILIKGIFSPDLFLKTYNNYLRISYXPELKKQKQLFPTHKQIED
 FWNCEVDKEHDLITISPHAEELTYTYREKTLNLNGLDNDKPHSPYTVTFSSIFEER
 PYHEGYLFFPPLSLSFQCNVPVQIHGTIHGVCNENLNTSIDRPSLKKTKTLLGSLPETSS
 QGKOLLERYVALVALVMSQHLSSDSALIKLTSFYTKENHHPFSDPEGLRKVLVEVYHLM
 SPQCIPLLSFLCWKTLDDCEKHFQAVLSAISEEANKNPSLPFWQFHNRIIEILNHV GAS
 ERLKILSLFRPGCEAV

0n_0738 830719 833895
 recB-Exodeoxyribonuclease V, Beta"
 KYFLFCEVVPKFNIFDSNSIQGKFFLEASAGTCKTFTIEQIVLHLEIGSLTHVHAL
 AITFTNASHTKRIKRIKDLAQTLPEKAVLNQZASPTFTLDINCNVQIKVMQRNALA
 LLDMSGFTIHGFCNVEIQYFPKPTLHKNPALHTSGLVHHITNYLKQDLKQNVLFQE
 FQHLAVRYNITKSHFTSGVLDKLAQTQPIQYFSGSRVERLQELGWHQIYNSLTFIE
 FQKFLDQLTAFQFQKQPTSLDLDHIFVDDLTSETHISLSLSPFKIAETFNKHLRAR
 YKRKAATVLENMGWRTLEKFLNLDRIPTNLVLQVLEQLQKQNTWPLSPDESVALEKL
 SYSSQAQPVQALREYQVLDIFDFTDQKQVLEFCHLISPKFTGSLTIDPGQSYI
 SWR:ADLLPTLTARSPEDKDLGLVNYKRTFKMEAIIHQEQKIFQCFIPLCLPGYLP
 IEYIAQKQSETPENPDPHVFHFFFTYTKDQALVGFTEALPQKQELPLNMNVLDGSDN
 DAVELIYATIPVIFPKNKKIPHTFTPHILTLLEALHTPENYKIKLIFLSPGLGL
 FEPFTKKEPDIYIQLARYIETIGILATFTPTMTQRIYLFSLSPGDIETQMEKLCGY
 IPTISYVPHLLHILKNEPFTPWEMHIAETGDELTLEKIDITTHSKKLECYDIPFCG
 EKKKNKSGSELLREMYVATPPAPYQVLYLTTPPGTIPGSAALINYVKLETCQSSAY
 AIIHQHGPIDPLEYSLPKDKHATFTPLFLLETTFLALTPPPTIPSPSTKFLDLTIK
 QDTPYKILTKQQLPDKETFTLHKFTFTIDFGLLDTQYELMTMRFIKHTHLBQ
 EFTLKLKLSKTFPTLPSQSTPQVHILIPFTFTPLFENQIHLNWFIDLFPEHE
 KYKYIDWKRTFLQFTNIDYKCHILITIKQFLDYQGLGVKAVKCIQVQIDVLEL

GVIFIPGUTUXNGFFALNGSEEDINFNPKAIKQCAVYH

CpN_0739 834392 833861
 CT 368 hypothetical protein
 CKVLFKLMGSLRNKTKICVYIIIALGILSFRSIPQEVYDKIRGSFVSLHVKFFPKIKQ
 APSGHLANLELNLVLRKAVGLEEKLEKLYEVSNHTPPLFPEILTPYFHLKLVGKVVYRD
 YTHWSSCWVNVKTHGKIKNSPVLGQNVLVGLVDVGVGHQSRIRLITDVGMPKPSVAMR
 QDIOGWWTKHCLRELIPQVEQISHAYILEKDYKESIQLOGLDGLIQGEGENALLRGIL
 LQVAKLWFLKLAIVHITDITKTKLHGLLTDGLVYTFKGLIAPYKVAPEED
 WMTFFHMLKLAELMELALALILMTRKTHHEDDIFVLLML

CpN_0740 836054 834864
 tyrB-Aromatic AA Aminotransferase
 SYMSFFNHITPSPDAIILGLQNVFFADKRPEKVLNVIGVYEHQPKRYGGLSCIRKAQTVI
 LEEBQNKSYLPISGLQIFLDEMRELVFGAVDPSAIVGQSLGGTGALHLGARLLSVARGS
 GKVVYPEQTWSNHIRIFSQEGLEVIRYPYYSKEQQLLFEPLIAFLKEVEKNVILLHGC
 CHNPTGVDFPTDMWKEALILMKERELIPFDTAYQCFAGHIELDRKPIEIFISEGNTVLV
 AASSKNFALYGERVGYFAVHTPTDELKIHISFLEEKIRGEYSSPQRWQVEITILSN
 PYLEKEWQSELANFIRESLGKMRTRFVQALRKVAGHTFDLLSQHGFFAYPGFSDKQVLF
 REQHAVVYTAGGRMNLGITEKNDHVVQSFQAYEL

CpN_0741 838383 836185
 greA-Transcription Elongation Factor
 EYIFRLTKGDIVDYLEKQVLEIEGQSANFLSLDEEYCFNDVVRGRELVEILEKVKSSSL
 ASLFGKIVDTVPLWEKIEPKGDKDRVLQLILDLQTSNQMFQDIATEVNVKYSCEENF
 NEALRVVGLRDGRDFQSLSRDFLMMHMKGVNFVHQQGQWGVGMVSFLQKVLIEFE
 GIMSAXDISFETAFKSLTPLSGDHFLSRRFGPDGFEAFAPKENPIEVVEILLRDLGPKTA
 KEIKDELVDLVEADWNRWQSAKTKIKKGTRIISPDNPKPEYVLSADGCSHMQLERK
 LGLSLNSAEKISLIYHFIKRLHSELKNIERKSLVKALQDLVEEGNKSLLQRELLLE
 YLGIKDAISDEKYEITSLSEDDTSRLLENMIVALQKSFSLVSRKYSFVQVVFQIILLYT
 TSPTRMDFVYKTIKNDPSSVEVLKRLDLSAHQPMFPELFWVFFLKLGNHEDGLDPED
 KEVLRFLFLESALNFMYQVASTPHKELGKKLHHYLVGQRYLAVRQMEGASLPFLKELLML
 STKCPQFSSDNLVQLSAEVQPTLKKHKSNEEENLVNLTSESFSRMKALQSLVGKE
 MVDNAKEIEDARSGLDRESEYKFALEKARLQEEIRVLSEENRARIILTKDLVFTDKV
 GVCGKVTLLKGDAEVEYITLPGWDADPDSCILSLQSKLAQNLGKKLNDVVLQKEYK
 ISRIQSIWEHGA

CpN_0742 838442 838888
 CT635 hypothetical protein
 TKMMVIMNSKSAKIIIDSIKQILTYINDFDPSFGSSLSDDADYEYLITKTEKIQE
 LDKRAQELITGTMSKEQMEVANNPNDPSPWEALAEKVRSSCEYRKETENLINEITL
 DGEFTKESKRPKQKLSSTKKNKKWIFL

CpN_0743 838956 840362
 "ngaA-Ubiquinone Oxidoreductase, Alpha"
 IEMKIVTNRLGDLDSLQSPKESGFYKNIIDPEFVSIDLRPFQPLSLKLKVEQDAVCSGAP
 IAEYKHFNTYITSHVSGVVTAIRGNKRSLLDVIKKTGPTSTEYTYDQLTSLSDLS
 EEFKENGFLALIKORPFDIPAITPTCTPRDVFINLADNRPTPSPEKHLALFDSMREGFVY
 FVAGVRAIAKFLGLRPHIVFRDLRLTPQELKTALHTLVGSGFPSPGSPSIHHSVAPIT
 NEKEVFTLSDQVLTIGHLFLKRLIHEQITAGHTALKSLRRYVITTKGASFSLLIN
 LNDISDNTLISDQPLTGRCLCKEEEPFLGRDHSISVLHNPTKRELFSLRGCFNKPPTF
 TKTYLQGFKKRTYINPDTNLHGETRPIIDTIDYKVMPIRIPVPLIKAVITKNFDLA
 NESLQFLVCGEDFALPTLIDPSKTEMLITKESLIEYAKESGILTPHOD

CpN_0744 841387 840389
 hemB-Porphobilinogen Synthase
 EMSSLTLSRRPRNRKTAIRDLAETHLSPKDLIAPFFVKYGNKIEEIPSLPGVFRWS
 LDLLLEIERLCTYGLRVMFLPTIIPDDLKDAYGYSNSNPKNILCHSIEHKNAFPHCL
 ISDALTDPYTHGHGDIPLNGEVLNDESIRIIGNIATLHAEMGADIVAPSDMDGRIGYI
 RSKLQDSQSGYKTSIMSYSVKYASCLSPFDRDLSSHVTSGGKQYQMNPNVLEALLLESS
 LDEEBGADILMVKPAGLYLDVYIRIGRQNTCLPLAAQVQSGEYAMLSAFQCGWLDKETLF
 HESLIAIKRAGADMIISYSAPFILELLHQGFEE

CpN_0745 841903 841742
 No robust homolog present in Genebank/EMBL as of 11/7/98
 VDSEFDWRASSLQGSTTYNAVDPKHTLAYGFCNQVSVKKPHLKPKPSQEKFL

CpN_0746 841939 843567
 CT632 hypothetical protein
 FSGRCPSFEVFMKGEEFTCKQKQCLSHFVNTLSDVFAKKNLPEVVKGALFSKYRS
 VLGLRALLKEFLSNEEDGDVCEAYDFETDVQKAADFQYQVLDNFGDDSVGELGGAHLA
 MENVSILAAKVLEDARIGGSPLEKSTRYVYFDQKVRGEYLYYRDPILMTSAFKDMFLGTC
 DFLFDYTSALIPQVRAYFEKLYPKDSKTPASAYATSLRAKVLDCIRGLLPAATLTNLGFF
 GNGRFWQNLIHKLQGHNLAEFLRLGDESILTELMKVIIPSFVSRAPHHHHQAMQYRRAL
 KEQLKGLAEQATFSEEMSSSPVQLVYGDPDGIYKVAAGFLFPYSNRSLTDLIDYCKMP
 HEDLVQILESSVSARENRRHKSPRGLECEVFGFDILADFGAYRDLQRHRTLQERQLST
 HHGVNFPVELLDTPEKSYREAMEANETYNIEVQFPEEAQYVMFMAYNIRWFFHVNAR
 ALQWICELRSQPGHQNYRTIATGLVREVVKFNPMYELFFKFVDYSDIDLGRLNQEMRKE
 PTT

CpN_0747 843949 844053
 CT631 hypothetical protein
 RTCMGCKGAQEVQILSSRLSQMKILSSSLFYKFKC

CpN_0748 844996 844121
 ispA-Geranyl Transtransferase
 GTLVLHALDTRPSIESAIEKALEGFGPIGHPIRSPVEYALQGGCKRLRPGVLCMMAQGL
 CLNHVMDSGALAEVFTSTLIADDLPCMDNDRPGRPTVHKAFDEATALLAGYALIPA
 AYSHLRNLAKLKEQCDPRCIDIAYNIIGDITDKNIQSGVGLQCYDDMFFGNRQGEHV
 QSMIKKTGSLFEIACISGWLFGGDDPQFAPITTFGNNGPGLLFQIKDDFDGLQKDSQOI
 GINYLALLFGEKAALELLARGQNNKLELLDRLSAGCLKNCESEFETIISGLQSF

CpN_0749 845638 845006
 qtmD UDP-GlcNAc 4-epimerase
 VCYMTYIACISLPHEDFIYHEIIRKAYITWDLDMQMLENHVFGHIGTVEGVTLKN
 IEKIEIAEPAYVEYAYIVGQILGQVTLVHRIALCPQNVITNIRVWGHCTEIKNSYLG
 HHTKAAHFAYIAEVLSEVNLAAVVRANRGLDGHPIYVPSDKKKIDTGRRLKQAF
 IAKGVAIAVNVVINAQHILPHTRIPYQVI

CpN_0750 846465 845707

tctD/cpxR-MTH Transcriptional Regulatory Protein + Receiver
 Domain
 KITDFILRIHSYNLFCFHMIGDKIILFVTEDELSLSQLKDLASQRSDYQILVSPVFTSF
 ESVAIFCEYLLLEQIFSPGIFPEEDLIVLPTFQEEAITKVLNQGATGYLLRPITAKVL
 DAVIRAFRLQHEVLEHSIPDTMTFGDHTFRVLNLVIESPEGSVYLTPEAGILKLLINR
 GHCLKRLKLLAEIKGNTKEIARNVDVHTASLRKKLGPYGSKIIVTIRGVGYLFSDDSDIP
 LQNHNDTAHPNEE

CpN_0751 846465 845707
 CT631 hypothetical protein
 MFRCILFGILLLTCFSSCGVLYYLFCSHDFSIGPKEKSRVWLEEEKEFTDSVLHHLPSQ
 HQHLHLILCFQGLLQKQKQFSAEKIFSKVYDEAQQDGPFLKEEILGSLINSFFLEKTD
 VMETILCLLNQRCNSPYHLFKALVCYKQKLYREVIEQLAYWQEEKTRALAPLLNISIE
 QLTDPLLDYISAHSLEIQKMFPEGRVILNRNINRLKHECEWNAKTYDRIATILLSRYF
 LELVESKADYFDYEMVLFYKKIYILEQCPYAEELLPEELVSLIMEHVILPKDKLY
 PLIQLEMMQKHYVHPNSSLVQILVDRFSTHMEGARFCEALVSFSGLEELHQIITTF
 EELLNKKVQKITEEAKQCVALLHLIDPSISISEKLALSDDTLQNVISGDDDEQHTKRVN
 LDLEAIIQSYVDIDRQQLVHHLVYGAKDLWKKGNDEKALNLLQLVLRFTIECESVVF
 LFIQAYKQALSSHAIRLLKLEKFISEANIPISIVISEAEKANFLADAEYLAHEDYDKC
 YLYSMMLTKVAPSPQSYRLAGLCLMENKRYDEALEFLCMLSPNDSINDYKTQKALAFQCK
 HQSKDRAAS

CpN_0752 848595 850082
 "recD-Exodeoxyribonuclease V, Alpha"
 GWAHTEFAFPLEDLVHQQVISPDLIAFASKHISDSFEESFVFLAVSSALLWRYGHPFLSL
 EENRIRPSLGGISETDLYRGFHNLPKEARDKLFWVVSGRLYRSYITRSKLLDKLSLCL
 SATPNFPSPSIDSSILSEQNFIENKIQGCFPSIVSGSGTGKTFLLAQILHLVQKPK
 LIAIVSPYTGKATSHIRQLMKYNIIDDMVLMQTVHVFHQEYAYRRYNSIDVLLDEGSM
 VTFDLLEYQVDTIQGYEKDKKLYTSSLIIIGDTPHLPPIIGVGNPLQDLIGYFHEMTTF
 LKTHRAKTVGVVQDLTQSVLRGEMISFSPSPSSIAIEVLKNNRFSKLSRQSEARLCLVTFP
 MRHSGVGLVNLNMTIHQRARSDDPLRIPIIMVTSRYETWGLFNGDGLLCKLTKLHFQ
 HEPIDSRALSOYVYVMSVHKSQSGSEYDEVIVIIPKGESEVFGVSILYTAITRAKRYRSV
 WGDPEHLIKIKKSNY

CpN_0753 851009 850161
 No robust homolog present in Genebank/EMBL as of 11/7/98
 IMATAHLGRQALLHLRSWTPAIRASGNLFRQQSMNLHNNVLFGADIVGAKINSTASRHA
 LGSSHYAAALQTEGFLGAADGVNTAVAGAMLWGQLLNGSMIFETDEETGLRRCNEAD
 AEGCMTQKLQRRSALTITGKVARLASKTLGTATFLHEMDVSVLSGANANKIGCKVTSCLNL
 VATGCSLTSSISLYRILSTRPETISDPENRNKPSAEFAARSKAIRNAFIAWLGDVVDLV
 CDALGTLSLFLPALIGVHAVLIMAILGLISCVINFKVYAKIG

CpN_0754 851381 851040
 rs20-S20 Ribosomal Protein
 QFILNLKVLVSGDIMAPKPKNKQVIRRRPSAEKRLTAQKRELINHSFKSKVKTIVKK
 FEASLKLDDTQATLSNLQSVSVVDKAVKRGIFKDNKAARIKSKATLKVNARAS

CpN_0755 851579 852799
 CT616 hypothetical protein
 YKDLFFMLLVKRWLHTCFKVIWYFLPVVTLTLLPLVCYFPFLSISQKIYGYFVFTTISSLGW
 FFALRRRENQLKTAAVQLLQTKIRKLTENNEGLRQIRESLKEHQESAOQLQSQKLNKS
 LFLHQLGLLVKTKGEGQKLETLLHRTTEENRCLKMQVDSLQCEGCKTEEVQTLNRELAE
 LAYQALNDVEYQATFSQORNMLDKRQIYIGKLENVQDLYMEIRNLQLESIDIAENISQ
 ESNALVTGNTISQLSSELKKAFAKAENIEAASSLTASRYLHTDTSVHNSLECRQLFDSLR
 ENLNGFLVYARQSORAVANALFKTWTCYCAEDFLKFGSDIVISGGQWMDLHSSREE
 CSGRILVTKSRGHLFPFRYCLMALNGLPCYHVLGVLYPLHKEVLQS

CpN_0756 852889 854676
 rpoD-RNA Polymerase Sigma-66
 ISYLPITKLSSKARNPLVLFQVRKLFMNTONSQATEVSSEESQKKLEELVALAKEQGI
 TYEEINEILPMSFTDPEQIDOVLIIFTGMDIQVNLQIDVERQKEKKKEALEGLARRTE
 GTPDDBVRMYLKEMGTVPFLLTREEVEISKRIEKAQVQIERIILFRYSKAESAISIAHYL
 ISGKERFDKIISEKEVEDKTHFLKLLPLITLLKEEDTYHLENNLLSLQKPDLSKQEAAL
 NDSEKCRITQAYLRCHFCHRNVTEDFGEVYKAYDSFLHLEQQINDLKVRNKRFAA
 AKLAARLKYKREVAAGRTLEEFKQDVRMLQRWMDKSQAEAKEMVSNLRALVTSIAKYY
 TNGLLSLDLIQEGNMGLMKAVEKFEYRRGYKSTYATWIRQAVTRAIAAQARTIRIPV
 HMETINKVLGAKKLMETGKEPTPEELAEELGLTPDRVREIYKIAQHPIISLOAEVGE
 SESSFGDFLEDTAVESPAEATGYMLKDKMKVEVLTLDRERFVLIHRFGLDLCKPKTLE
 EVGSANVTRERIRQIEAKALRMRHPIRSKQLRAFLDLLEEEKTGTSKVSLKSK

CpN_0757 854709 855134
 folX-Dihydroneopterin Aldolase
 PKIKNALVIAERYQLIISKFRMNLFLGCSVEERHFQPVLSVTFSYNEVPSACLSDK
 LSDACCYLEVLSLIEEIANTKPYALIEHLNLELFDLSVISFGDKASKIDLEVERKPPV
 NLLNPICKTISKELCPSVLSA

CpN_0758 855104 856459
 folP-Dihydropteroate Synthase
 RAMSEPRFVCLSLQSNLGNRFKNLQIARTLLGQAVLGRSSVILETEALLPGSPPEWD
 LPYFNSVLVGETLLGLRELLVTIKQIEKVVGRAEESPPWSPTIDVDILYGDSEFCCDH
 TEITIPVSNLLSRPFLIALIASLCPYRRFCTQGSYPHNFTFGLAHLHLPSPGMIRRSLS
 PDMTLMGVVNTNDQMSDGMFLDPEKAVAAQAEKLPTEGAVIDFGAQTNPVKQFLSV
 DQWERLEPVLRLKLETWSNRKQYPIISLDTFYPEIILRAMDIPIQWINDVSGQSSMA
 EVARCELSLVMMHSSSLPVDPKNILSFSVPFGEQLLSWGEKQLKMFSDVGLNANQVIFD
 PGIGFGKAAQSLATLYEIAKFKRLGCPILIGHSRKSFSLFGNHDPKDRDWTGVLISL
 LQQQGVYDLRVHNAAHQKALSVAACEACAPI

CpN_0759 856434 856997
 folA-Dihydrofolate Reductase
 LLVKKVPIHNFENFLGVENCKNPPVPGIYACDPRGVIGLEGLKFWHYEDLQFFSETIQK
 FPVIMGRKTMLEPKYFVDRAVVFCEKRLXVHGEIWTSLSEFLLLDLSSPTFIQGG
 GELYCLLENQIVPDDFISHIYKAYAGTFFELSLLELDTWTKTVLRDQVITTCYENHHS
 QNTKNIEL

CpN_0760 856997 857694
 CT611 hypothetical protein
 RHGPKLECLIPKPKQVPMKHTTPTKTHYIYFYDLYGLESLLPKLNFPPIVITTSKIVS
 DLPGAVVEKCVIPDELIRQADAYVFEKYIYLPKKGILIPGAGIDENVRKVFVLY
 PRFLLLSVNTLQGLWLNKYLHLCGIIITGHTPTPLRPTMGLAIAWNPFLVYVWKP
 DCPGRALKMTYNNLLDGLSAAAVIAWLYDEQHTPIAIEEAFKITPISPTTQUMTLA
 IAEDELYGLDGLQMAWETPAIT

CpN_0761 857498 858375
Ct610 hypothetical protein
GIMTSWIELLDKQIEDQHKHEFYQWSECKLEKQQLQAYAKDYVLIHKAFFPCYLSALH
ARCDLQIRRIQLENLMDDEEAGNPNHIDLWRQFALSGLVSEELANHEFSQAQACMVATF
RRLCDMPQLAVGLGALYTYEIQIPQVCVERIRGLKEYFGVSARGYAYFTVHQEADIKHAS
EKKEMQLTQVORENPDAVLQGSQEVLDLWNLSSFINSTEPSCSK

CpN_0762 857498 858375
Ct611 hypothetical protein
BELLVFSRFSEKQNYGLONMETKRSIYMLNDRPKKALEAAVAYIEKQFGAGSIMSLGRHS
ATHEISTIKTCAISLDLALGHHGVKGRVIEIFPPESSGKTTLATHIVANAQKMGVAAV
IDAEHALDPSYASLIGVNIIDDLMSQPDGCEALSAELLARSAGVDVVIDSVAALVPK
SELEGGIDGVHVGLOARMSQALRLKATLSRQSCAVFINQIREKIGVSGFNPTTGG
RALKFYSISRLDIRRIGSIKGSNDSDIGNRIKVKVAKNKLAPPFRIAEFDILFNEGIISSA
GCILDLAVEYNIIEKKGSWFNQYQKKLGGREFVREELKRNRLFEEIEKRIYDVIAANK
TPSVHANETPQEVPAQTVEA

CpN_0763 860520 859972
ygfA-Formyltetrahydrofolate Cyclo-ligase
NFMPTDPKIEKSLRKLKISIRDLSEERKHEASSAVASVRSFSKESVLSFVSNFHEI
DMQEAENRILIQKCTLALPKIDQENLYPVLISPDILISVHPDQPFSGKTPISSDKITHV
LVPGLAFDQGYRLGYGHGFYDRWLQAHPYPSIRTIGIGYCEQKIDRLQESHDIPLSQI
YLC

CpN_0764 861819 860524
Ct648 hypothetical protein
GYKSMIDIKKFLCFLCSLISAMSPYIGKTDGYEKLTLTGINIIDRNLSETIKSKEKLK
YTKVDFLAPQYQKVMRYKXKRGDNVSCLTAYHTNGQIKQYLECLNRRAYGRYREWHVN
GNTKIQAEVIGGIADLHPSAESGWLFDQTFAYNDEGLEAAIVYEKGLLBSGSSVYYHTN
GNWKECPYHKGVPQGFYTYTSSGKLLKEQNYQGGKRRHLSIRYSEDSSEEDVLAWEEYH
EGRLKAEYLDPTHEIYATIEHGNGIAIYGYKAVIETRAFYRGEYGYKTRFDNSGTQ
IVOTYNLQGAHGEFFFPYETGPKPLLLNWHGILNGIVKTYWPGTLESCKELVNNK
KSGLLTIYYPQGMATEEYDNDLLIKGEYFPGDRHPYSKIDRGGTAVFFSAGTITK
KIPYQDGKPLLN

CpN_0765 862415 861801
Ct647 hypothetical protein
TTIYIKLLGRMLKWIILISLFLSLLSILPVLAITINHVKISQRWSDNLSQILTLKVIR
DHEDQVIRKHNRIKSDRNNLSIESLNASCQRLPSKERERLNKLSNLSLAQSKEWER
KRLAESNHLQVWVNEQMHNDFAVRLEQATEMDNEDIESLFLNPNPVPAPLVFTFCW
KMTKQTPPLGNEVWLTHAEISRWI

CpN_0766 863785 862394
Ct646 hypothetical protein
ANNEKLPYHHTIGLTKAENNTIKAILQKTKGWIHVCEQIEPEGKTWSLKKYFAAPTTF
SLGGSDILVSSSSSLKRNKRNILKVALTNLEASLALPWSLIVQPLGKPTDRGETPLTL
WIAQNTLKKELSLFSLQAQIFPDKLSCRAADIFLAEASPLKSLPAYLLIYGGESEVTCI
FVQAHIAVARSFNNHSTKSCDDIATLQIQIETFPQTPLPAIHVAQISPNLQKILEQK
LSLPLVVCQSMTYGVEDDEWEIYGTDIAAAHGASRRPLTFYDTSVSPAAQKHWLLRS
SLIIGKALMATVVSLSVSLKLSLSSASNNHFAFACPEESVPLRSLLKAAETKAIQK
KNSAENYPLPTPTSEQLTKFLALGKSSPSIKFSYFSTMTSPYSPKDNPSLPYSALVE
VKGQSPEDIPQFLKKISSHPKLQHVSESLEDQRSFKLQFTLSS

CpN_0767 863878 864177
Ct645 hypothetical protein
NMLSYLLRTAINVYSFLILAYIFASWVPDQCSARWYQVLSKCDVPLNFFRRFVPRIGF
IDPSPFVGLLCLGLPFFVILRVLRFIILNIFHSFWLLQYL

CpN_0768 864144 865163
yohI/nir3-predicted oxidoreductase
YFSPMAAIFIKNILLRSSIYVAPLAGFSIDYPRYMSALYQPLMFCMVVKGILYAP
ERTSKLLDYNMRPIGAQLCGSNPETSGEAAKILEGLFDLIDLNCQCPDKITKIDGSG
SGLKTPELIGRILDKIINSVSIPTVKIRSGWDMEHINVEDTVRIIRDAGASAVFVHGR
TRAGYHGCQKQEIYISRAKAAAGKEFPVFGNDIFSPEAAQAMLTTCGQGVLVARGTGA
PWIGKQIQDYLTGTSYEKIPFIKRAAFLEHMRVLDVYQSETKFLSETRKLCGHYLSA
AKYFRLRSSLAKATSYQEVYQLVNDYEADDSLETFFVKC

CpN_0769 867763 865121
topA-DNA Topoisomerase I-Fused to SWI Domain
SIQGPAAIRLMKKSLLIIVESPAKIKTLQKLLGSEFVASSIGHIVDLPKEFGIDVDHDF
EPQYQVLPDQKEVINIRKLAACEKVLSPDPDRGEAIAWHIANQLPDSPLIQRVSN
AITKNAVTEALKHPRITIDMALVNAQAARRLLDRIVGYKISPLSRKLQQRSGISAGRVQS
VALKLVDREKAIDAFVPEVYNLRVLMDQPKTTKTFWAHLAYAQKKWKEIPEGKTEN
DVLLINSEKARHAYAELEKSSYTITRVEAKARRFAPPPFITSTLQOEASRHFPSASR
TMSIAQTLPEGVLDSEDSSTGLITYMRTDSVRVDEALTTVREYIQOTFGKEYLPEKANV
YTTKMTQDAHEAIRPTDINLTPOKLKNSDDQDFKVVNLIWKRFVASQITPAIYDTLAV
QITTDTEIDLRASGSLKFKGLFVAYEEKODDQEDDHLPLPLHAQDALIKEEVSEQE
AFTKPLPRFTTEASLVKELEKSGIQRPSYATIMNKIQSREYTTKENQRLPTELGKIISQ
FLETNPRIMIDGFTALMEDELELIADNKKPWKLLQEFWTTPLPVVITAEKEAVIPRIL
TNIECSKCHKGLVKIWSKNSFYCYGSEYPCDYRTSEELAFNKEDYADTPWDSPCPL
CGGVKMKVRHGRYCTFLGCEKYPCGRGTISIHKKGEEIEQEPIPCPAIGCNGKIPKRRSR
YKIFYSCEYPCSVIGNSIDAVITYKSOTEIKPYKKKTPTKKSSAKTTKAAKTPSKK
GKAKSSVKSSSEKKTGPLFLPSPLAKMIQNEPVSREATKSIWDYIKEHQLAQAPENKKL
LVDPNNLAIIGPNPIDMFQSLKHLQHLQHTLVKYNDESSASS

CpN_0770 868322 869131
Ct642 hypothetical protein
KPTRNVEKLEFVTCLSPDDDLITFNKQCLLAGPEEEKVAFVLRSNAMLDAGPETAASF
PEGLREQDFIPPEYVEVLYSNEGLDWEACQTLNINJEVTLQRKHHRKASPLWLMQYSRD
E/LAHEAVIAVARMKHEPVEEVLAVQTSRWQRWRRFPGLPERSGESYLLLEFTTLGLGI
SLWYPAQILIMLVLMYMLRLCMAQSYLYRAMKKIPXMLGVPPVLVLRDLTKEIKMFA
KEIPVLEHYPARKKLENVKWKQIQYQSYFV

CpN_0771 870511 869144
rpoD RNA Polymerase Sigma 54
IPFGNCKELVDKALDMFQKQKLSIKYPLRLRMOGQLMLQSPLETLSYYVDEIINDP
FFGLSGLEEEWQINRYPTNITFYLNQTVNQESLTPLLQIEEAFUAEERFIAHQI
AYHLDGHLFLNIPEDFAQELPLLEKTHKVNQTLQNLQEGTASPSLQSYVMKLLRNS
HLCAYGVKPDYPLMTNKEPAPIMKKFSLDCELRNLLKALQZIPWCPAAACTVPMVS
TFLDLYLFGSGGWKTEVTRGLPSIKLNKTFHFYCHLKEEQNLISQILSAKWLK

NLRKRECTLLQVMTLLPKQEDFLLGKIPAPYPLGIKDLAEDLJFHEOTIFRAIENKAVA
APGIFIGLKHLPFGIHDSSHSKENVLOWIRONTATEQTFPLSDGVISDRITAKGIPCAR
RTVAKYRAQLKILPANKPKKLFYIRSSNSHFRDRQF

CpN_0772 872400 870469
uvrD-DNA Helicase
KLGLINTCISELNEAQRKAVTAPLNPVLVLGAGAGKTRVVTYRILHLINQGIAPREILA
VTPTNKAAERLEKPIVNOCASTNEFDVPMVCTFHSLOVFILRRSINLLNRENNTIYDQS
DALNTHKALNTHKALNTHKALNTHKALNTHKALNTHKALNTHKALNTHKALNTHKALNTHK
LDRANLTHKALNTHKALNTHKALNTHKALNTHKALNTHKALNTHKALNTHKALNTHKALNTHK
NVFAVGDPDQSIYJWGANIHNILNFENDYFNAKVLCLEENYKSYGNILNANALIKMNA
SRLEKELRSVKGCGEIRLFLGSDTDEEADFAAEILQLHRVGNIKLRDICTFYRTNSQS
RTFEDALLRRIRPYEIIIGLSFYKRKEIQDILAFIRIFISKSDIVAFDFTVLPKRGIGS
TIFATLQYAIAGQLPILKACQALDQDKVLSKKQEQGLQEYALFPQIEHAYNTLSLR
DFIESVVRITGYLEIKEDADTFKDRKSNLEELYHKALESESNPKTHLELFDLDDLALKG
SDDDLNLADRVNMLTLHNGKLEFRVSPFLVGLGEQLLPHANSLQGTYENIEERRLCYV
GTTRAQDQLLYLTAQVRSVWGTVMRMPKSRFLKEIPKDYMIQVR

CpN_0773 872485 873195
ung-Uracil DNA Glycosylase
FMQNAIDQLPVSWOQLPLCWRQQLKEEWSKPYMQQLIFLQKEYKEHTVYPEENCVFS
ALRSTPFDQVRVILGODPYGKQQAHLGSFSPVEGQRLPPSLINIIFREKLTDLGLENHK
GCLQSWANQGLILLNTVLTVRAGEPFSHAGKGWELTDAIVTKLQERTHIIFVLWAGAA
RKKCELLFNSKHQHAVLSSPHSPPLAAHGRFFGCSHFSKINYLNLNKNPKMINWKL

CpN_0774 873183 873425
Ct606.1 hypothetical protein
LEAMPNBEIHVSVCQKTPRLTAKSVSMEMLLTQQLPSAEGMPSVANLEADFLRAEALL
AEMREIRGCLEQSLRTLVPSE

CpN_0775 874040 873414
yggV family
FERMKIVYASSSHGYKIRETKTKLRLGDFDIFSLSDFPDYKLPOEGDSITANALTKGIH
AANHLGCWVIADTMLRVPALNGLPGLSANFAGVAGYDKHRKLLDLMSLSLSDVRS
AYFECCVLLVSPNQEIKFYTGICBGYISHQEGKSSGFGYDPIFVKYDYQTFAELSEDEVK
NQVSHRAKALQKLAPHLQSLFEKHLTRD

CpN_0776 874180 875487
Ct605 hypothetical protein
FIFVLKNFYDCLLMFFQFLSFTMKKIFYSFVLLSCIFPYVGAQVFGVGLDRIFSEGEYTR
CIQCKKLLISHSAAINSRGQDALSIFYSRKHDCVTEILCTLEHGYGATPTTETVGNOPS
RYPNRLSVSLYGVKEVPKEVAEHCDFVVDQDQIGVRSVFTVLMQIVASADLKGQFI
VLDRPNPMGGRIVDGLPNTTSGSLAIPYCYGMTPEGLALFFKKTYPANANVVIIMKG
WNRSMTFDETLGIWMTSPQMPDPQSPFFYATGILGALSVASIGVGYTLFPKVLGAPWM
DGEKVADELNRMKLPGVLFLEFFYEPFFGKYKMEMCSGVLLVLQDPKIFYPVETQCTIWG
VLKALPKQVEQTLKIERIPARRSSICNLFGGDEFLSISHKERYIWPRLRLCKESRES
FHQLRSSCLLSEYAES

CpN_0777 875586 877178
groEL_2-heat shock protein-60
TSEDVWVFKSQFEGSLALKRGVHALTKAVTAPFAGPRGYNVVIKKGKAPIVLTQNGIRI
AKEIILQDAFESLGVKLAKEALLKVVEQTDGSGTALVVIDALFTQGLKGIAGLDPOEI
KAGILLSSVMYVQQLQQAIELOSPKDVHVMVAANHDVLTGTVVATVIGSADLKGQFI
SKDSGSIKTRGLGKRVKSGVLSYFVTRPETMDVWEEALVLILSHSLVLSSEELIRYLE
LISEQNTFDETLGIWMTSPQMPDPQSPFFYATGILGALSVASIGVGYTLFPKVLGAPWM
ATLIGQESSENCEIPVSLDVLGRVKQVMITKETFTFLEGGGDAEIIQARKQELCLAIARST
SESECELEERLAIFIGSIPOVQITADTTQORERQFOLSRATKAMKGGIVPVGGV
AFLLAAHAIEVPANLSSGMTFGFETLLQAVRTPLKVLQAQNCGRSSEEVHTILSHENPRF
GYNGMTDTEFEDLVAGICDPLIVTSSSLKCAVSVSCLLTSSFFISSRTKT

CpN_0778 877400 878092
tsa/ahpC-Thio-specific Antioxidant (TSA) Peroxidase
APVAGSDRVPGYERGGQRFESSLRVNRNKRVEEVEFMTLSLVGKEAPDFVAQVAVNETCT
VLSKDLGKGVVLFYFKDFTYVCPTELHAFGLGEFHTRGAEVIGCSVDDIATHQWL
ALKKGGGIGETIYPLLSDEDKVIRSRYHLKPEEELSPRGVFLIDKGGIIRHLVNDLP
LGRSIEEELRTLDAIFFETNGLVCPANWHEGERAMAPNEEGLQNYFGTID

CpN_0779 878502 878095
Ct602 hypothetical protein
RFDLIFQMKFTVALFGAEKGSYDTAYFCRSLVDLHNYLGDVSSPGITLAKTLLSDYNV
YFVRHREBYGVDSYFFGLHFLNTQTTLKNIIAIGLPGVGNQHIIEASRSLCQKHNLLL
FFDHLVDLLTTFNQPF

CpN_0780 879241 878591
papQ/amiB-N-Acetyluramoyl-L-Ala Amidase
HONKIAVQSLRFMHAKLSFFILLSLLSGIDCSRLHAAGRSPSLQGVLAIEDISAKLAS
HEVEIVMSERLDEQSKQKQWTAAPETLAQKIRELESQKALAKTLAVLTTSVKDLQT
NLQSKLQEIQKDRALAQDLRLVRRSLALVDSSSPGAYADPSDPVPIYIVREGDSLS
KIACKYKLSVELKINKLSDAIYAGORCLQPNQK

CpN_0781 879851 879198
pal-Peptidoglycan-Associated Lipoprotein
QNCYRSRMTYVPLGCFPSATDKENTMNIHSLWKLTLALLALPACSLSPNYGWEDSCN
TCHTRRRKKPSSFGVPLYTEEDFNPNTFGEYDSKEEKQYKSSQVAAFRNITFATDSYT
IKGEENLAILNLVHYMKKNPKATLYEGHTDEFGAASYNHALGARRANAIEKHLRQGI
SADRSTISYKGEHPLNSGHNELAWQNNRTEFKIHAR

CpN_0782 881077 879773
fobB polysaccharide transporter
GIDMLRQCLQVFFEFASLVYAEELVVRTEHITLIEVSCQTDPKDKIKQYLSL
TEIFPILAIADLQNTAARKESSPLAISLRHVPQLS/LLQSKRTPTQTLGPTIQON
LIVDQYTHHAADTVIYALTGIEGASAGKIVFALSLSGLDQKLKGQELWTTDTPKRLAP
ETITYGLSTPKWVWVNFYLVGYPYGVVIFLGLSEHTEGKVLPLKRNIMITIS
PKYLLALVAITATNPLIQPESLTDGMPQEPRLLEHTEGTQNPAPNPEGLVLEIS
NKLALFLYIMLDPFQALRLTKKYPNGSCPAWGPDPKIAFGVLEKVRQICITYDIS
QDITGLPTPTNKEHWAIDSRHIVFGAGNAEESLEYLISLVTKTKNKAIGVCHKRP
PWAFAPQPIKRL

CpN_0783 881088 881109
Ct604 hypothetical protein
IMMKYLYIATATKATLILLVPAEPLPKRKLQPKAFQELVITIQPKPIVITPFWVIV

GYEFNELGNINFGTLLLLNLNIEKADIDYHTEKRLQKELNLLVADLALLPDPFTFPKV
TFSQHLRPRQLSGHFNFTVQDXQDGLTGILGLAGIDGLPJDALRGLRLFLAYSSD
VSLQISKDTADTSGFKTTEGNEAVVMTFIKYVEKRLLELLSLGEGATMFLQGESFV
RLPLETHQALQPDRLICLTGTEADILKYSFQLIEELLKQPLNPLNTENLIDSLTMLNN
ETFSADGTLITLSEF

CPM_0793 896838 894919
 rbsb-sigma regulatory family protein-PP2C phosphatase (Rsbw)
 NLTQLVPLNVDVLSGLPSDVLDDLDIAPEETPNVLLSNEMQVFGIYNEISLKIYFPGND
 KYVASSIPEHLGNNHKKIDIPKNTFFLAALQSPKNQEVFSVMQANVFDAKTOELQGI
 LITTVSASESLKDLLINKHSYLTAKTILSKGYVILKASDPALHLHTVYEDMTKEKFCQV
 FLNDDPCPIDSELGLTSLPLDTIGVEYSFKIKDTEIWGCIENVPSIDIAVLSYAKKES
 FAPLWRRRARYTAYFFCILLTGLSLIAFIVARRLLSLIRKLAATNIESRKNKQVLYTDDSLG
 FEINRLGIFNAMVENLHKQHLAKTNEEMKENAQNALLHGEQAQRLPLNTLPSYPHIE
 LAKAYIPATVGGDFVDFVYVGGESKARLFLIVADASGKNAGCYSGLTKLNNMLTFLSR
 SSLSQAQIGETSLRFLYNNKSGMFLTLCVYCHTNTMEYVSCGHPACVLPDGETS
 WLFHSGMAGLGLPEVANTISKLPHHPKGSGLVYSDGITEAHNNNNNDMPGEERLQAAIQG
 LTPGSDADVHRLMLSVKTFVGNSHDDITLLILKLVES

CPrn_0794 897123 898004
No robust homolog present in Genebank/EMBL as of 11/7/98
KSSKHSRFLLLKSGGNVSLYQKWNMSQLKSLCYSTVAALIFMIPISQSFADSLIDLNL
GLDPSVECTGSLGAGFVSVPYTKAGSTPVEYQFFKYDSKTTFTLSAVETANGSQYAYGIS
YDGTIVTGLCSLDGKAVNGAKWSADGTLPLTLTGTSHTAEARSISKDTQYIEGFSYDA
SNGPQAVGWASGATGTVQLADISGGSSRSYAYATSDGDTIIVGSMESTITRKTAVKKWN
NVPIYVITLGGDASTGLYISGGDTIVGAANTATVINGNOESHAYMYKDNOMKD

CPn_0795 898008 899195
No robust homolog present in Genbank/EMBL as of 11/7/98
TLGGANSSATGVSSDGVIVGQAQADKSVHAFQYINGEMKDLGLTGGTSSSTAKTVSPD
KGVINGRSQIDGAWHAFCHTDFDSNNVLFDLNTQYLRENGROSLINFLNGMLQR
ASDHFEFTGFRSNLITAGCLYVNLQNLPSNLAAQYFGIAYIKRPKRYGLVFLDNHFSH
VPNNFVNSHNRLLMGAFIGQWDSGLGSSKVVSKYQKGQKATITREQLENTAEGSGESH
EGVAAGIEBGRYKSLGSHVYRQVFFGLQGFVHITRKETYENAVQFPVHYHDPIDYSTGVVYL
GIGSHLALVDSLHVTRMGMEQFAAHTDRFSGSIASIGNFVFEKLDLTRAFAMRNVN
YELPYQLNLILRVNQQLPQGVMGFSDSLTRALGFL

CpN_0796 899280 901340
No robust homolog present in Genbank/EMBL as of 11/7/98
SELYSSYLQPLCNMSTVIRNSALPCLSRSETPKKVSRHMKFMKVLTPWIYRKDLWVTF
LILTAIPGSAFTLLVDIAGEPRHAAQTGMSGDGKIVIGMKVDPDFPATITVGQYVIGDGLH
PLEAVRPGVSYVNGITIPDGTIVGTNAYIGSGVAVKGVNGKVSLEMLPDTLDSVASA
VSADGRVTGCGNRNINLGASVAKWEDDVITQLPSLDMNACVNGISDGSIGIVTMVDV
SWRNTAVGWIGDGLGVITGLGTTSVASAITDGTIVGVSGENADSGTHAYAKNGMVD
IGTLGGFFVSLAHAVSDSGSVIVGVSTNSEHRYHAFQYADGVMDLGTGGINATYAAQVGS
DGKVIIVGRAQVSGDWHLPCFPQSPAPSEHVGQSTVTVQNRGMDPNATYSYLKNSQ
QQLQRLLIQHSAKVESVSSGAPSTSVKGAISKQSPAVQNDVQKGTFLSYRSQVHGNVQ
QQLLTGFAMDKLASAPKCGFKVALHYGSDQALVERAALLPYTEQGLGSLVSGFGGQVQ
RYDFNLGELTVLQPFMGIIQVILHSGVASEYSEKNVRYPSVDYSVASAATFSFMGHVFA
SLSPKMSATLGVTERDLNLSIDEFKGSVASMGNFVLENSTVSVLRPFASLAMYDVRRQQLV
TLSVVMNOOPLTGTLSLVSGSYNLSF

CpN_0797 901552 902694
No robust homolog present in GenBank/EMBL as of 11/7/98
VLILTVINVLTKLGLNMSKKIKVLGHTLCTLFRFGVLCAAALSNIQYASTSQESPQKSI
EDWKGYTPTDLLESLGSEEAHVA/GNGSRISVAGSAGGQGVSTAVIWEHLKHLGLTG
GEASSAEGISLQGEDEVVGVMSDTREGYTHAFVFDGRMDKDLTGLATYSVARGVSGDGS
IVGVSATAGEDGYGVQGVKWEKKIKLQKLKLPGQLVSEANAISDGTGVTVGRGEISRNHI
VAWKVNKNNAVYSLGTGGVSASAEAI SANGKIVVGVSTTNGGTHAFMHKMDTHMDLGLT
GGGFSVATGVSGADRAIVLFGFSVAKTGEIHAFYYAEGEMEDLTLGGGEEARVFDISSEGD
IGTSKTLDTAGAERAYLVFIHK

Cp_n_0798 902810 903856
No robust homolog present in Genbank/EMBL as of 11/7/98
VVFEIIFVVRVPMKTKTCQNVYSIGVVFVSVLFEVLDTLAFGHIDITSGGLYSWARGV
SGDGRVVVYEGYEGNAGFVYDGEKLEGLLEGLVRSEALVFKASVYDGSVIGITSDQDPSRAV
KWINGALVDLIGIFSEMGSGFAEGVSSDGKTVGCVLSDDTTEVPAKWWDETGMVVLNPLP
EDRHSCADWDASDEDSGVIVGDAMGSEEIAKAVYWKDGEQHLHLLSNIFGRKSSAHASGKDS
FTVYGEISSENEVHAYVYHNGIKIDITGLGGDYSVATGVSRDGVIVGHSRTTRTDGEYRAF
YKVDGRMIDLGTIGGSAFSGVSDGDKTIVGKTELTGECGHAFIYLDL

Cfn_0799 905001 903940
No robust homolog present in Genebank/EMBL as of 11/7/98
KRENNNAIAKQILRRSLMSQSSLMVMVLSLYLSGTYGVITDKPEDDPHSSAAVKWDHWGK
TTLRLSNKKASAKAVSGTGATGVTGFKIDTWSYTAAYRVNKKTPDPTLSWKKWSKATG
ISDDGDSIAGIIVENELQSSFAVTKWKNNEVLPTTWAGVSKAYGISDDGVSIVGSAKDAW
TRFPAVTKTQHEAQLPCVMAWAGVANSVSGANSIGVGSVDGASGILVAKWENGTITHL
TDLQCYSAITAKGVSNNNGKTVIGRSESTVYGEVFAHPCKNG/MSDGLDLCGYSAAKGVSA
TKGVYGMSTTANGKHLAFKPYVGGRRMIDLGEYSWKEACAAHVSIDGELIIVGSGE

Cp_n_0800 906550 905249
 EnoLase
 RKKEIKMEFAVDIAQIREILDSRGYPTLHVK/TTSTGT/GEARVPGGASTGKEALEFR
 DTDPPRYGQKGVLAQVKNVKEILFLPVKGVGV/EQSLDLSLMMDSQSGPKNETLKGAAILZ
 VJSLATLHMAAAATLRRLPRLYLGOCFACSLPCEMMNLIT/IMHADDNLFGEQFPMTRIGA
 SIKEAVNMAGAD/FFRLLKLLHRLGLSTVGDEE/GFAPILASNEEALLLELLALEKAGFT
 PGKDLGALDA/AAISFYNVKKVTD/DPHYHEGIALSLNLD/DRYP/IT/EDGLEAEDYDQW
 ALLTEVDEKXVO/AAIDFLPTNPEELLLEGISHLNAN/VL/IKTNQ/IGMTTPTVVA/IKLQAM
 AGYTT/ITLIR/NTTPTT/ADLAVAFNAGQ/IKTGQSLSP/EPVAKYNIP/MEITREELGSEAI
 TGTGCVN/ND/DEIT

PTN_0901 208109 206727
 VTP EKXINAEKAP ARS ARSNTB B
 TPTTTLPLIAIAIAAQVQEAAPLACAGVQV/KSQV/LPTN TPTTTLIAIVANVIL
 PLVLAHNKTLAAQIAAQVFRFFENAVVEVFTZDYDQVPEAYLAPDITLPLALLINDE
 LKDKPLATRETLERKKPLVLVAVGCTYGGCENXTMALVLEVGVINPINTLTAQVVK
 THQVA TPIIDPQARVAVLVIDPVEELALPLFLPLTLTITZDPLNITLKEVP
 ATATVPEVPEHYVTPENRQQAQPTGGLLILPRAFFADDPPEKDRLEPHTHTLMLKEIG
 PCGLINERHPTHTPNTNIAQNTLALVLEDEPDLITDCHPTLTPHAMVKGQAKKQGL

CLEELVDFHIESEALPCHTFPDFGCVFTEGFPPEECTIPESPPQAPKTLKIALIGRP
NVRKQIINILLNEERCIDNTDPTDHDILYSHKDRQYFIDTAGLQPKVSKNSIE
WISGSRTEAIPRADICLLVLDATQKLSYCKRILSLISKRKKPHIILINKWDLSEVRM
EHYCKDLRATDPYLGQAQMLCISATTKRNKKIFJAIDELHHVVSNKVPPIVNTKLASA
LHRNHQVQCGRLRIYYAIOKTTTLPQLFLFINKASLLTKHYEYLLKNTLKSSFNLYGI
PFDFLEKPKPKRHN

CpN_0845 958151 95850

CLP-X-CLP Protease ATPase
RPLEDIDIANASPTIVSTIFPDVIGIGVAFGIIIVVQDGRLEFVATGSGEYKDGHRP
DRIFSSMREDALRRDFTVNGMYDFFEDKVFDFVEGTRDIEKKVIRAIHGPRLRFSDEK
LRLRAIRFSSSLGPTLDPTTERRAIKEAPALVNSVSFERIWOELKMKLRQPYGALSLL
LKRLKVLIFIPPELRDIPYSLRTTIEFARKNPTPLFEILFLPLFGVSEEAATVAFGR
LRIINSELKLIESWYALPHFQNSGNRVFAWHLASPTAPLFLFLSALQKPSRQOHF
ISRVQELSRLEQFILRIKTSSPVVAPDLIAKGISPRGLLDLLREAEILSIENECLDK
EKTLLLLQKQGFWK

CpN_0846 959383 958112

CLP-X-CLP Protease ATPase
REHMKNNKLITCSCGRSEKDVKEKLIAGPSVYICDYCIKLCSGILDKKPSSTISSAPVSE
TPSQPSDLRLVLTPEIKKHIDEVYIGQERAKGELVAVVNYHYKRIRALLHNKQVSYGKSN
VLLGLPTGSGTDLIAKTLAKILDPVFTIADATLTVEAGYVGEDVENILRLQAADYDVA
RAERGIYYIDEIKIGRTTANVSITRDVSGEGVQALLKIVECTTANVPKGGKRRHNEQ
YIRVNTENILFIVGAGFVNLDKI IAKRLGKTTICGSDQADLSQKTRDHLAKVETEDLI
AFGMIPFVGFNCTVNCESLDELVAILEPTNAIVKQYMEFABENKLVFKKEALY
AIAKKAKQAKTGARALGMILENLPRVDFEIPSDPTVEAHIQEDTIAENKAPITIRRTF
EAIA

CpN_0847 960019 959387

CLP-X-CLP Protease Subunit
KLFDDEOTMTLVVYVVEDTGRGERAMDIYRLLKDRIVMIGQEITEPLANTVIAQLFLM
SEDPKKDILFINSPOGYITAGLAIYDTIRFLQCGDVNTYICIGQAASMGALLSAGTKGR
HALPHSRMMIHQPSGGIIGTSADIQLQAAEILTLKXHLANILSECTGQPVKEIIEDSERD
FPMGAEEAISYGLDKVVTSAKETNKDTSST

CpN_0848 961556 960177

CLP-X-CLP Protease Subunit
VQASSPAFPKSNKKGLVPRSLNEDGFSVDLEESPGCISALVKSVEVLNKLKQALK
KIKKEITLPGFRKGAPODDVIASRYPTNVRKELGELVQGLVGLDGRPLSPKAVR
SNSATQDLQEGAKVEFSYEAFAISDLFWENLSLQEEAASEISDSIEKGLTINIGMF
ATKRPVPERPSQEGDFISLHVSKSNDENASAAIFENKYFKLSEEMTDAFKEKFLFS
TGHVAVETITSPETQSFLRGDTLTFTVNAVIEVSIPEIDDEKARQLAESLDDLKAKRLI
QLEKQAKDKQKRFSEADALAMLVDFELPTSLLEKRLISLITREKLLNARLISQSEDE
LEKRSSELKEAEEDATKALKLLFLTHKFIKDEKLITRELEQYMMVCSRERFQOQPK
DISNLTQELVMSARDRLTYSKAEHVLKALLASTPSEA

CpN_0849 961752 965285

CLP-X-CLP Protease Subunit
mot1/snf-SWF/SNF family helicase
ADYHHSYSGEMLNFRKLRRDFSANILQDGKLFQEGAVIDA KILSMNGTVCISAOVR
GLYDNITYECIEVDRSESDTVDSNCCDSYNYDCQHIIVALLFYLEQYFNEMVAVARSADL
ETDHEINEEVKELKETFAAATKEEREDREHKEILREYVHAANALSANFFFLPLEYL
EKDSAEALAVLFVSVNEDTPAPANQPIEFQVLVRLPCRSKPFYISNRTFLEGLVQEPYV
LNGRRFFPTMQSFNASDRKLIDLLIRYVRYPNHTTEKRLKSAYLMPALPGLVIAKMFHF
QLADRGGSLGKESFSGLFCGNLEELPCWSLTPAKMKFNLDFFDMPYKALLMTPTILVD
DDEVQPEQTMLESADPGI IHHFVYHRFSQKRAHLRSFRLRDIAIPEALFGSFRENA
LPVEQYAEIANVHLLNSFVTLPRVDEVAICDSYLDGELKHLFLVGLSRVPAASLA
LYQYQDVRAFISDEGLARNLVEERKMLEEVFGSFIYDERDAGFRVSKSEKIVFETETIP
ANQHRTIFNCPENLSGQFIYDETIFELSFPSGGTINNYEADLKGVLKGVLLDWDICI
SAKRFLELPKAGQOSKGTTRGVNSGKLPCILVLDLEKIAPVVQIFNEIGFVLRLDQV
KCPWLSLTGISLDQFEALPVNFSMERLIEIQOQIRGEIEFDDQVPOQIATRLSYOTE
GVHLERLKRMHLNGILADDMLGKTLQAI IAVTOSKLSQSGCSLIVCPTSLVYNNIE
FRKNEPEFRTLVIDGVPSORRKLQVADAEVTSYNLQKQVLYKSFRFDDVYVLDDEA
HHIKNRTTRNAKSVMIQSDHRLITLTGTIENSLEELWSLFDLMPGLLSSYDRFVGKYI
RTGHEMGNKADNMVALKKVSPFILRMKEDVLDLPPVSEILYHCHLTSEKQELQSYA
ASAKGELSLRVKQEGFERIHIHVLATLTLRKQICCHPAIFAKDAEPGDSAKYDMLMDLL
SSLVDSGHKTIVVSQYTKMLQI IKKDLSESGIFPVYLDGSKNRLDLVNQNEDEPSLLVF
LISLKAQCTGLNLVAGADTVIHYDMNNAVENQATDRVHRIGQSRVSSYKLVLTANTIEE
KILTQNRKSKSLVKVINSDDDEVSKLTWEVLELLQI

CpN_0850 965254 966390

CLP-X-CLP Protease Subunit
mreB-Rod Shape Protein-Sugar Kinase
LGKYYVNCRRYDFMSPHRNLFLKLNFSNRLYNRALGRFDKVFNFSSGNGVGLDGTANTLV
YVRGRGIVLSEPSVAVDAQTHAVLAUGHAKAKMLGKTPRKIMAVRPMKDGVIADFEIAE
GMLKALIKRVTSPRSVFRPRILIAVPSGITGVKRAVEDSALHAGAQEVILIEEPMAAAI
GVDLPVPHAEASMIIDIGGGTTEIAIISLGGIVESRSLRIAGDEFDECIINMYRRTYNLM
IGPRTAEIKITIGSAYPLGDQLEMEVGRGDQVAGLPITKRINSVEIRECLAETIQOII
ECVRLTLEKCPPELSADLVERGMVLACGGALIKGLDKALSKNTGLSVITAPHPLAVCLG
TKKALEHLDPQFKRRKGNLV

CpN_0851 966378 968195

CLP-X-CLP Protease Subunit
pckA-Phosphoenolpyruvate Carboxykinase
REFGIVMWNSTNIKHEGLKSWIDEVAKLTFFKDIRLCDGSDTEYDELCTLMESTGMTIRL
NPEFHNCPFLVRSSADVARVEQPTICTSTEAEGPTNNWRDPQEMRRELHQLFRGQCG
QRTLYIVFPCMGPLDPSFIVGVELTDSYVVCMSKIMTRMGDDVLRSLGTSGKFLKCLH
JVGKPLSGEADVSWPCNPKSMRIVHFQDDSSVMSFGSYGGNALLGKKCVALLRASMA
KQGWLAHEHMLIIGITNPEGKKYFSAFSPACGKTHLAMLPLPGWKIECIGDDIAWI
RPGRDGRLYAVNPEYGFVGAPGTSERTNPNALATCRSNSIPTNVALTADGVVWEGUTE
QPPEPLTDWLCKPWKPGGSPAHPNSRFTAPLRQCSLDPEWNSPQGVPLDAIIFGGRS
ETIPLVYFALWEHCVTIGACMSGTTTAAIVGQLKGLPHDPFAMLPCTGYNMAYFOHNL
FPAENRCLKLPKIFGVNWRKNNQJETLWPGFENLP/LWEIFORTDQLEIAERTPIGY
LHIIQKFNILKILNLDLQTVQELFGVAGWLAIVENIYDYLKIFGDCPQIITDELLRIK
DEIKKK

CpN_0852 968274 970113

CLP-X-CLP Protease Subunit
CT1711 hypothetical protein
IKRLIDYLYIINTVTLQPSYINFTNVTALSOGLDTGAIIDELKALFPQCLQDKAQ
IKHALHJWELJAEALKAQVOTISGLYLPTESSSRFISGAIIDRTMPTTDEVKAILQ
NINFTPTKIFVBLDKVTKYGLDSVTPPEGIDENPELALINUYITLLNKLKPKFAAGST
IPHIADYHALYALIDFVKIEALKKADAPPKKVAHFQWIMTYNNMQVLSYPTDYNL
VQIAHJLNTAAQVQYKKNFYGLERDLIRWTDQATHYPADAEYNAPDAGYIQSL

LNLSCNYRQLTENMLPNTTSLPOEIAQIRSFQNVNCTIAGNLTLPMTMLDTLGGV
ITYYQCCATIPGMSYGTSTPAKQNTIDAINGEASQWQARANGFDVTSQVFDQFATNIQS
GTSYRQCDLIFQNMKNVEINPFLQQAASFLRYPNYLMRSRMYQTIEDAANRSITALDGLI
SGWSTQIATFTQTKNSLDPSLLKYFDTMKANKESFVTAPLQMVYSSMLLDKYLPTQQNV
IASLGIMTYSNKAARYLHELKEITTFQSAIDIYYSLSYLLQNMNLQAVADPIGKAVGV
NDEKTRAMADITCRNKIKAAIDKMLVEIKADAESKSSQIRELVDTLTNFKSQSDDLIRNL
SCLLGLFSLGLTLKAVNDPNATYEAFTAEIETPEFNKNKRLATFESFVYIQGGQNGITPGG
QOOLQAMESSQODSTFTIIONQALQLLESSAMQOEWTLVSAALALLNQMVSKIAARRIKS

CpN_0853

CT1712 hypothetical protein
NIMHPKIEKNSLPLTAVAPVEESYHPSVATTVDYVDATLSRLHTVLKDVKEARNLD
LGKAFITSMKQGFINTGTLEAIQASLADQSSRESRKKEKIFHQHGLKAAPOAATATSG
VQPTADPVMKPLQSAFAYVLLDKYIPAQEEALYALGRENLNLSGAYQNLFSPLDMDIKS
FNSAPINYNLGSYISQTSQANFAYGYEMILSRYNNEVSQCRRLDIASVAKAALANMSA
SVKANVSLTDAQKKQIEDIASYTKSLDVHTQTLTDVMTNLASTITFVPLNLYDPSYRIV
GGDLISIALQNDKVLVDGKVDITAVNEGQGLNFTTTLVTDVQNYGDLAQTLQMLDLE
LKAMQOQSVLSASLKLNGMYTTVISGFIN

CpN_0854 972849 971806

CLP-X-CLP Protease Subunit
ompB-Outer Membrane Protein B
GFFDMNSKMLKHLRLATLSFMSFGIVSPVAYVALGAGNPAAPVLPVNPQGTGWCACFL
CNSYDLFAALAGSLKFGFYGDYVFESEAHITNVPIVITSVTSSTGTPTITITSTTKNVDFD
LNSSISSSCVFATIALQVETSPAAILLDIAFTARVGLKQYVILPLNAYRDTSPNPLNA
ESEVTDGLIEVQSDYGVWGLSLQKVLQSGVSVGVSDYRHGSSPINIYIVYKANPE
TYFADTDGNLSYKWSASIGISTYLDNYVLPYASVSGTNRKAPSDSTFELEKQFTNFK
FKIRKATNFRVNFVFCGTTCISNNFYYSVEGRWQYRAINITSGLQF

CpN_0855 974001 972994

CLP-X-CLP Protease Subunit
gpdA-Glycerol-3-P Dehydrogenase
GLMKQHIQVYLGMIWGFCLASLLANKGYPPVAVWSRNPDLIKQLEERRHPLAPNVISPN
LSFTTMDKAIHNAFMIVEGVTSAGIRPVAEQLLQITDLSVFPVITSKGIEQNTGLLLSE
IMLEVLDGVSPTPYLGLSGSIAKEVNLGSPCVSVSAVSDYTKQIHEAFSLPTFRVRY
NTDILKGAALGGALKNVIAIACGIAEGLSGFNNAKAGIVTGLHEMRKLAINDCKPETLN
GLAGLDLCTVCTFESSRNLRFGLHLLAQGLTFEQAKAKIGMVVEGAYTALSAYQVAKHHK
IDMPTITGIYRVLYENLDLKEGIALLLQRTKEEFL

CpN_0856 975410 973995

CLP-X-CLP Protease Subunit
Agx-1 Homolog-UDP-Glucose Pyrophosphorylase
GSRDRNVRLVTMTESVYSPSAMHVNLSADLKAINQHEILDIPSLSPKQQRLLFQQLTS
VDIDFFRKQQLLSSPTAILKDFHPTISFASSEDPERAHAGTTLLEKKEKACVAVLAGGQ
GSRLLCKPQKGLFPVSPIKKKPLFQVLAEKVRAASKLAGQLPLAFMTSPNTRQTSRFF
ESNDYFHLDPNQVDFCQPLWPLLTLSGDLFLEMDTLALGPNNGCITATLLYTSGWHEK
WKNAGIEMVSIPIDNPLALPFDVLCGFHMSNNEVTIKAAALROTIEDVGLIVKSHVS
GKTSVIEYSEIPONERFALNEDGKLYCLANIGLYCLSDMDFIRHAAYQQLPLYVVKHAK
QVTKVLEKSNLNAWKEEFIFDLFCYSOHCQTLVYPRQECFAPLNKLEGNHSPDVTVRQALS
DRERQLPHKVTGKLSPTNTFELEADFYYPSTSTSLHWNKAFFEEPFEAS

CpN_0857 975808 975392

CLP-X-CLP Protease Subunit
CT1716 hypothetical protein
LLLLRQYIKTARGISRLMRDLGSLSLIKVKIHKYLDTLHNQKRLATVSRNIQATNKR
IADLHREYEHFISRDNIKHIDILEYKTLQSSLYKQSSERLFLTHHQQLQELINNR
KIIEKIKNNYKSDQIEGT

CpN_0858 977115 975757

CLP-X-CLP Protease Subunit
fliI-Flagellum-specific ATP Synthase
RNSETNRNRRTRPSTFCFDSMNLNKEKLIHNNWQPYRACGLLSKSVGNLIEVDGLSACL
GELCKISSTKDPNLLAEVIGFHNHTLLMSLSPHLSVALGTEVLPLRRPSPHLSOHLGL
RVLDAGFNPIDKEDLPKTRKPLLSLPPSPMMRPIDQIFPTGKAIDAFLLTGKQRI
GVFSEPTSGKSLSLALGSKSTINVIATIGERGSRVREYIEKHSNALQOORTIIIAAP
AHETAPTVYAGRAAMTIAEYFREGQEVILFIMDSLRWIAALQVLAARGETLSAQHAYA
ASVTHHVSSEPTERAGNMDKGSITALLYLYPKHPDIFTDYLSKLTDGHHFLLTSQKALA
SPPIDILSSLSRQAALALPHHYAAERLRLSLKVYNEALDIHLAGYATPGQDEELDQAV
KLLPSIKAFLAQLSSCYLNTLKLQLEALDS

CpN_0859 977597 977055

CLP-X-CLP Protease Subunit
CT1718 hypothetical protein
VFLVTTQSPGSLSSQSLPHPHDPWDTEPTSLPEDPNKASQELHSLVHLFRKLSIHLS
EVEKTVQQLKPLDLEALLICEKFLYKLENQPELALLSLTALQRTHTLRLSLTPKVLFLH
PEDKTLTDWISTHELPMIKHAEFFPDTSCRRSGKFIETPNGLRQBISEELDHLVSLA
A

CpN_0860 978639 977608

CLP-X-CLP Protease Subunit
fliI-Flagellar M-Ring Protein
RTLVEFQNLAKLTALGISPLGCLLIGQVVSACILFGRRSSNPLAPTOVKTEKSGNWLK
LTQMNKPLIESLTKEQLEKDLTSFHTASAKVAIALSTEDDMSPLHLSVILTRKEE
SLTPSLFSTIDVLCSSFLKREHISLSDNGLNLYIPESITVNSLFIHTLENLQKIFP
KEHFALYHAKAEKPTLTLNENYIAHLTKESEKIVAHTKHYLYQNYDDSDYVIELT
FARLQNHKSFPKVLIGSMILVISLMI/ALASFLYLAHAYERVSEPRKIRKGINISKL
LEITQKESPEKIALILSLDPPKAEALLNRLPEDLKHOVLKYKL

CpN_0861 979752 978925

CLP-X-CLP Protease Subunit
nifU-NifU-related protein
ASYPTWKFLMTLPLEPMIFWSGLSAK/MKKFLTTPHCACTFSEEDAEAKEHLVTKGQGH
PLMNCNVTFYWLVDKKNVILDAKFQYFGHPYLIPLAEAVCNLVCKGSYSEAYKMTLDDI
DKSLRVHAHQALPEDSISLYHVFIDALDTAVEQCELEPLEDGSPLQNSPMNLDDEFAN
PYSQSDWEALTEKLYALRATIAEKIGPYIAMDQGEVTVESLENFIVTIAYSGNCSOCP
SSLGSTLNSICQLLRAYTYFELQVYDESSLNLCHP

CpN_0862 980124 979722

CLP-X-CLP Protease Subunit
yfhO-NifH-related protein
GRGTIFPITDGTATLMEKIONKRAKPFILWLMHQVAIPPSERVKEISYALHSDIPSLPPG
TAKLAKEETESIRQLWALDIIIFRF/PHFHT/VIIVLAAVLNGLMPCQRNIIILIPAH
DQQLLINSICRHHQGLTATVWTVNHEVIEVEGLIETLSPRLSLSLAAHGLTVIOP
LDLPLCLCKDRILLIILDIILIGRAFLPTEILADIITFSSAALGMSIGITIRKSL
ERVFCWFPPTSAHAFVAAQMTAFER/IALPIPTFHTNIAKKLIQELQVLDPSI
QLAFSEVQNRLENVIAVAFDIPAEILAFHILQAGYPLQVYERTQPLAQVQLNGIGDPP
LGLHGLHSLTERKQLEI/KIARAMHIAIKHLPTEGLGSS

CpN_0863 981520 981520 980831

gpmA-Phosphoglycerate Kinase

113

114

CpN_0928 1058009 1058557
Thioredoxin Disulfide Isomerase
CHHTQQTYYTRFPKDSDMKFWLQCAAFVCCLLTLPCCAARRRASGENLQOTRPIAAANL
QWESYAEALHESKQDHKPICLPFTGSDWCMWIKMQDQILQSSSEFKHAGVHLHMEVDF
PQKHQPEEQKQKQELKAQYVTCGPELVFIDAEGLQARMGFEPGGGAAYVSKVKSAL
KLR

CpN_0927 1059818 1058670
*THLPS 43 kDa protein homolog_3
FIISILFLPLALLWLKTKCQFFILPSSIIISQSMKTAVAIRRMFLSHIKQLLSKEI
SAADRIVYQYDDLVDSLAIKIPHALPHRWILYSQSGSLMENLFDGDSLSHLQAKATG
SNLLVFNYPGIMSSKGEAKRENLVKSQYACVRYLRDEETGPKANQIIAFGYSLGTSVQAA
ALDRVETDGSQDTSWIVKDRGRPSLADVANQICKPIASAIILVGMNIDSVKPSERLCR
PEIFYNSNHQDELISDGLFERENCVATPFLLEPEVKTSGTKIPERDILLHLNPLSPNV
VORLAIVSYNLDSENKRSQDD

CpN_0928 1061035 1059884
*CHLPS 43 kDa protein homolog_3
RRKDAFTLLNLSNRSDILSGISNPHVSVFSSTHAKQLSDFSCKKHPILTIVTIIVKI
FKLLIGLIIPLGLIYWLQCLVSLALFPRSSMLYSVLTKCFKYRLEQEIQDYFVNKLDP
SFKDPAVSEKRTIQDHLITDLAIHFSTARPKNWILSLGSGDFLEMDIGLKDSFLR
SWKELAKLLGANILYINYPGVKSSGCKLNLLENLAHNLCAKYLQDKIQGGPANEIITYG
YSLGQGVQSAALQKNPFTNSTSWAVKDRAPHSLLPAANSFPGPIGLTAVLARWMDA
EKNSRELPCPEILVYSADRFPRSEVGDDTALLPEPTLAHAIKRTPFARSKKFGEVNLH
SSPLKHPITQKLAAEILSLSRKN

CpN_0929 1062301 1061186
*CHLPS 43 kDa protein homolog_4
EKFMAPHGNAFVEDILHSHSPQATYFSSSTRAQKLHEFKDRHVLTRIASVVIKIFKV
LIGLIIPLGLIYWLQCLVSLALFPRSSMLYSVLTKCFKYRLEQEIQDYFVNKLDP
SMRRVPIQDNVLDITLEICLSQAPTNRWMLISLSDCSLEIEACKEIFDSWOFKALIG
ANILVYINPGVMSSTGSSSLKDAHNICTRYLKDKQEGQGAKEIITYGSLGLIOAE
ALRQDKIVANDDTTIAVAKDRCLPFIISPEGFHSCRRIGKLVARLFGWTKAVERSQDLPC
LEIFLYPTDLSLRSTRVONKLLAPETLAHAIKNSPYQNKKEFIEVRLSSDIDPIDSKTR
VALATPILKKLS

CpN_0930 1062851 1063330
No robust homolog present in Genebank/EMBL as of 11/7/98
NKMSLAPCSTGLQVPHQVHHALDTRVITLTAACLSLIAGTVLVLGAAAILPSLFG
VIGMILHILFSSIALIYLYKKTREVDQIALEPLPMISKQDSIIDFVKTRDYASLEKKAT
FAYTHYLYDGSVMFYREIPRFLMGLKDLAKRDMRQALF

CpN_0931 1064078 1065718
lysS-Lysyl tRNA Synthetase
IDERYLQWKSIDIYTNILEERMTARAEDLHEDFLYRSHKLQELSELGVVLYPYEFPVGS
CEDIKKTFASQELGNSAAMSRSTPRVRFAGRLVLFAMGKNAPGQILDHNTIQVMFNR
EFTSYHGLSEDAEITPIKFIKKLLDGLDILGIDYLFTHSGELTVLAVETVLLCKSLLS
LPPKHAGLSDEKVEYRKRWLDDISREVSDFVKSRYIILKILNMDAHGFLVETPTLQ
NIYGAEAKEPFTTMEALHSEMFLIELEIALKKILVGGAPRIYELGKVRNEDGIDRTH
PEFTMIEAAYMYDYKEMVFNENLVEHVRVAVNHDNTSLVSYWKHGQVEVDFKAPWR
MTKRESIATYAGIDVDHSDQKLEIKLKKTTFFPETAFATASRGLIAALFDELVSNDLI
APRHITDHPVETPLCKTLRSGDTAFVERFESFGKGLKCNAYSELNDPIRQRELEQOH
TKRELLPDESECHPIDEFLEALCQGMPPAGFGIGVDRVLMLTNAASIRDVLYFPVMRR
PDKEKTN

CpN_0932 1067160 1065721
cysS-Cysteineyl tRNA Synthetase
IDERYLQWKSIDIYTNILEERMTARAEDLHEDFLYRSHKLQELSELGVVLYPYEFPVGS
CEDIKKTFASQELGNSAAMSRSTPRVRFAGRLVLFAMGKNAPGQILDHNTIQVMFNR
EFTSYHGLSEDAEITPIKFIKKLLDGLDILGIDYLFTHSGELTVLAVETVLLCKSLLS
LPPKHAGLSDEKVEYRKRWLDDISREVSDFVKSRYIILKILNMDAHGFLVETPTLQ
NIYGAEAKEPFTTMEALHSEMFLIELEIALKKILVGGAPRIYELGKVRNEDGIDRTH
PEFTMIEAAYMYDYKEMVFNENLVEHVRVAVNHDNTSLVSYWKHGQVEVDFKAPWR
MTKRESIATYAGIDVDHSDQKLEIKLKKTTFFPETAFATASRGLIAALFDELVSNDLI
APRHITDHPVETPLCKTLRSGDTAFVERFESFGKGLKCNAYSELNDPIRQRELEQOH
TKRELLPDESECHPIDEFLEALCQGMPPAGFGIGVDRVLMLTNAASIRDVLYFPVMRR
PDKEKTN

CpN_0933 1067532 1068578
predicted disulfide bond isomerase
PVILLQNIKRCSLKQLKVLATLLLSLPTLEAAENRSDSIVWHLDYQALQKSKEAEL
PLLVIYSGSDWNGPCMKIRKEVLESPEFKIRVQGGKFCVCEVYELKHRPOVENTIQOONLAL
KSKFKINELPCMILLSHREERYIRIGSPFNETGNSLCHIVESDLSLLRAFPMTLS
SLSELQRYRLAEELSHKEFLKHALELQVRSDDYFLLSEKFRLLVEVGKMDSEECRIKK
RLNLDKPNKQKTHFTVALIEFQELAKRSRAGVRQDASQVIAPLESYISQFGQDQKDNLW
RVEMMIAQFYLDSDQWHAHQAEVAFEAAPNEVRSHISRSLEYIRHQ5

CpN_0934 1068948 1068526
rnpA-Ribonuclease P Protein Component
YFVHPLTLPKQSRVLKRRQFLYITRSGFCRGSQATFVPSRHPCTCRMGITVSKKFGK
AHERNSFKRVREVRHVRHQLPNCQIVVFPKGHKQPVFVSKLLQDFINQIPEGLHRLGK
TKATTGGECTPKSEKVTAPR

CpN_0935 1069100 1068957
r134-L34 Ribosomal Protein
EDTVKRTYQPSKRRKNSVGFTRMATRNGKLLNPPRRHGRHSLVDL

CpN_0936 1069330 1069470
r136-L36 Ribosomal Protein
YLMKVCSYKADPKQKDKLVRRKGLVYINKKDPNPQOQAGPARKK

CpN_0937 1069487 1069798
r14-S14 Ribosomal Protein
VKRMKKSVAKEARRRRLVEANFKKRDLRKIVKLTYYDEEKENARITLNKMKRDTSP
TILHNPGLTTPRRKYLKKAISRCIFQMAKMCIFVYKAW

CpN_0938 1070355 1069849
CT788 hypothetical protein (1100 aa) (49) peptide periplasmic
QFINLSTYTLMLPLSLIILYVILKCLAYIAOKKYPITIGWFFAIAEFGIGLVVLLLL
PDRHAKELKILDPDNCIDFOLKSLKAGNDLPLVGLDQIVLUTENWFLNKKDRENV
GPIIDFEEAVLILKQKYPITILWKKMKDKVKKVLPDQALAKFVAK

CpN_0939 1070629 1071175
CT790 hypothetical protein
HINRWITRLSLTLIISTVLVFFSEIEELIGGGKMEKQNLKLDVKEIEFPETVFSRDIETR
VIQVILHCLAKINGVSLLOGLNIDALFGRIERMKGIYVEQDSKNHLVKVRVENVVDYG
VSIPEKTEEIQQVISEISETYGLHVAHVHVIKGLTQPKDRIDEEIEEVSVDQLPSPE
DFLEENSEG

CpN_0940 1073030 1071204
CT792 hypothetical protein (1000 aa) (49) peptide periplasmic
YFVHPLTLPKQSRVLKRRQFLYITRSGFCRGSQATFVPSRHPCTCRMGITVSKKFGK
AHERNSFKRVREVRHVRHQLPNCQIVVFPKGHKQPVFVSKLLQDFINQIPEGLHRLGK
TKATTGGECTPKSEKVTAPR

CpN_0941 1075504 1073018
mutS-DNA Mismatch Repair
VMTKEKTPMMEQWHQCKEAGDSVLLFRMGDFYEAFFYDDAVLLSQHLELTLTQRGQIPM
SGIPVSTVDTVDRLIGKQFKVAEAFQGEPAKEKESKKGPMARDIQRVFTPGTLLSST
LLQKEKFNMYIVAINRIGSLFGFACLDLSTGSGFFIEECENTKELVDEICRLAPEVLSCKN
FYNKETAIVMQLQHLKLTSTYADWAFEHKFAQKLTTHFQVASLDGGLGLVPAINA
AGGLLSYIQDKLLLPKTHAIPTQTRGKQKLLIDTASQVNLLELLAPLNDPQKNSLRIM
DHTSTPMGGRLLRQILISFPYNPKEILVRQDAVEFIRQVTLRKNITFYLCQVDRIDERM
TKVTGLAGPRDITGLRDSFSAGAQIYEQLASATPEFFIDKCSLDTKLASLIALLSKL
NGDPLRVSQDNIIVDEFHNDLKLRLHNOEHSQEWIWEYQERIRKETGKKLICKFAQAL
GYIYEVSSSEFAPQLPKDFIRRSRLHAERFTTIELQGFODMNSIEKLTLETQFFKDL
CSHILQRLITLALSOQLADLDYIISLADLAHQAGCRPHVDMSDTLCTYRGCHPVNT
VDGTGIPNDTMRGSGQTRMILLTGNMAGKSTYIRQIALLVIMAGMGYSIPAKSAHIGV
IDKIFTRIGAGNLSKGMSTFMVEMETANILNHNATSLVILDEVGRGNTFYDGLAHRFA
AVVEYLLFTDKKAKTLFATHYKELTTLEDHCPHENFAGVKDKAGQPVFLYELKQHS
QKSPGHVARLAGFPLCVSRAQQILRQLEGPESTRPAQDKMQQLTLF

CpN_0942 1075955 1077754
dnaG/prim-DNA Primase
NCSITKLRTMYTEESLDNLRHSIDIVDLSEHILKRSATYKACCPFHTEKTPSFIVN
PAGAHYHFCGGAGHDAIGFLMQLHLYGSFTAEIILVSKKFQVDLVLPQKDSYTPPGQLK
EELRHINSEAEFTFRYCYLHLEARHALQYLYHRGFSPTDIRFHLGYGPGQSFLQAME
ERKISQEQQLHTPTGFTGNKWFLLFARRIIFPVHDLGHTIGFSARKFLENSGGKGVNTPT
PIFKSRILFLGNLSRRRIAKEKVLIVEGQADCLQMDSGFNCTVAAQGTAFTEEHVKE
LSKLGVKVLVILLFDSDEAGNKAALRVGDLCQTAQMSVYFCKLPQGHDPDSFLMORGSSGL
IALLQESQDYLTFIIEKMSYKPKFGPREKALLVEEAIRQIKHWSPLILVHKLQALSL
MMVPEDMLSLANPOVTAEPONIPKQVPIHPIHIVMETDILRCLMFCGNSNTKILITAQ
FYFVPEDEKHEPCKRLFAFMIYSEYKRYKRVFDEACQVLSQSLQLLTKRRLNTEALD
TIFVQSLQKMDRRWREQCKPLSLNQNIQDKKLEILEDYVQLRKDRITITLLDPESELIP

CpN_0943 1077972 1078238
CT794.1 hypothetical protein
FPMKSFPLPFLSVILCCGNLLSSPSRAISVITESIGMSAVKTLVLSEKAHEFLEGIGY
GVGASSILRDWQTCQWLEIESLLAGNEVM

CpN_0944 1078503 1078997
No robust homolog present in Genebank/EMBL as of 11/7/98
IKIMHRYFIPLLALLIFSPSLVRAELQPSNRKGGWPTQSCAEGSLFCFEAAVYNN
IEEGKPGILVFFSERPTPEFADLTNGSFLSTPIAKGFNVVLCPLGISPLDFHKKMDPV
ILYMGSLFLEMPFEVAVSGPRCLYLIDEQGAQCAVLPLETKN

CpN_0945 1079001 1079660
CT795 hypothetical protein
SIFKMKILPSYFGHNFQDLRRHYMRIALSLLSLMIFPIGEEESRPGSEDSNTQEIWG
SQDTQCLYHSYEQGLQASRIEGKPLVIVVLCNSGDDGQACTIGLSETCEEVLVLSGSI
FSELANFVVLVPSGVNPLIYPPIDPLAIEIVKFELFDESFPGLSITVVGVTPEGPG
DIEVSPVSLTVEEETLPSEQTTEVESTSELSQEDPAIA

CpN_0946 1082816 1079745
glyQ-Glycyl tRNA Synthetase
GECQCKKCYTLESFVSEHPLTLQSMIATILRFWSEGGCVIHOGYDLEVAGTFNPATFLR
ALGRPEPYKAYVEPSRRPDGRYGVHNNRLQNYHLQVILKVPENFSLSYTESLRAIGL
DLRHDHIFRTHDOWENPTIGAWGLGWEVWLNQMEITQLTFFQAGSKPLDTISGEITYGI
ERIAMYLQKKISYDVLDNDTLTYGQITQASEKAWSENFYDANTEMMFKHDFEDFAEAL
RLTKMLGNSPAYDVIKASHAFNILDARGTISVTERYIARIRQLTRVADSVVEWRAS
LNYPLLSSCTSEPKTESVVPVMSSTEDLLIGSEELPATFVFIGIQLESARQVL
TDHNTIYVEGLVGLSPRLALLVQNVAREVQKAFKKCPMLTSLSPDGDVSPQOQOQF
ASQGVDSHLYQDLRSRASHLARTVNGSEYLFLLHPEIRLRTADILMQELPLLIQRMKFK
FNVWNSGVVEYARPIRWLVLYGHEILPTLGTITASNSFQHRQLDPRKISISSPQDYV
ETLRQACQVVSQKERRMIIEQGLRAHSSDTISAILPLRLIEATFLSEHPFVSCQOQFSEQ
FCALPKELLIAEMVNHQYFPTHETSSGAINSNFFIVVCDNSPNDTIIEGNEKALPTRLTD
GEFLPKQDQLTPTTIEKLKSVTYFEALGSLYDKVERLKAHQVFTSTFSLAASEEDLDI
ATQYCKADLVSAVNEFPQLQIMGEYLYKHANLPTASAVAVGEHLRHITMQOKLSTIGT
LLSLDLRLDNLALACFILGLKPTSSHDPYALRRQSLEVTLVASRLPIDLASLLDRADH
FPSTIEEKVWDSKTIHEILEFIWGLKTFMGSLEFRKDEIAAVLIDSATKNPIELTOLA
EALQLLKDHEKTELAVITTHNRLLKILGSLKLSMTSSPIEVLGDRENSKQVLDARFQF
PKETSAHAFLEYFLSLADLSNDIQDFLNTVHIANDDGAIRNRLISLLLTAMDKFSLCHWE
GVAV

CpN_0947 1083433 1084059
g4S4-Glycerol 4-Phosphate/4-phosphatyltransferase
GSRVGLPNYITSRILEITPILMLYLGKWFYITPVVLYVLLALLAISLTDALDQVVA
PKFQV/TDLAKLLDMANIVYRIITYLTITTOPVNLPLLLYFIPLARDVYELTRTVCAF
KRVYVAAARAKKALITQVYFLLVLMVLPVHALLQONLEIFAVTVYITAVYITAS
GIEYFMNKNILQAKTKNDEPHHLLD

CpN_0948 1083483 1084047
g1qA-Glycerol Synthase
GEMPIYQVAVETPIVVRARADAVACERKILAKQNDVFLLDHYHILKIKKLSQVLEF
RQPYEFELGQQAASATVYVYGLTITITLDDQLELFTTIVYFNNVVRNFAAAAATA

YLQEADPADIVHLHWHVOLLAGLLKNPLNPVHSGIVFTIHNFGYRGYCTSTOLLAASQID
DFHLJHYQLFRDPQTSVLMMKGAALYQSDYITTVSLTYVQEIINDYSDYELHDAILARNSVF
SGINGIDEDVNNPKTDPAALAVQDASLSEDPVLTFRKEENRAVLKYLGISSDYFFLI
CIVISRIVEEGKPEFMKEIILHAMEHSYAFILIGTSQNEVLNFRNLQGLASSNIRLI
LDFNDPLARLTYAADMICIPSHREAGTQLIAMRYGTPLVRKQGLADTVIPGVNGF
TFPDTNMFNEFRAMLSNAVTTYRQEPDWWNLNIESGMLRASGLDAMAKHYVNLQYSLLS

CpN_0949 1085887 1086483
r112-L32 Ribosomal Protein
FTALAGLAVHNNIMLAAITPPTKYNLRLTGLHNNVYVWYVNHNTVDALNT
KKFLSLNESGALSSTVFSLSYEGRIKALVKDIOQYITTVYDHLDFEELVEDRPIKNI
PIRCINAVDCIGVKGSLRQVIRAVRVVCKPKDIVPFLDVRVSGLSQTRKLSKIKIP
AGIETITPLKEVAITVSR

CpN_0950 1086470 1087027
pH-Peptidyl tRNA Hydrolase
PSLEDNMKPLIVAGNPRHGYANTRHNAGFLADRLVEELQGPFFKPLSKCHALMTLVS
SSGFLVFKPTTFVNLGKAVVLAKYFNVALSHILVLADDVNRSGFKLRLCFNGSGGGH
NGLKSTASLGSNEYWQLRFGVGRPLEEGVELSNFVLGKPFSEENLQGSIFVEASTLFT
EWGSKF

CpN_0951 1087113 1087457
rs6-S6 Ribosomal Protein
EFLMGKNENQLYEGAYVFSVTLSSEARRKALDKVISGITNYGGEIHKHNDQGRKLAYTI
RGAREGYVFIYFSVSGAITEWKEYHLNEDLLRFMTLRADSVKEVLEFASLPE

CpN_0952 1087469 1087723
rsl8-S18 Ribosomal Protein
GENMKPVRHNEHRRKRFNKCCPFVSAGWKTIDYKDVETLKKFITERGKVLPRRITGVSS
RFQGVLSQAIKRARHLGLLPFGVED

CpN_0953 1087727 1088248
r19-L9 Ribosomal Protein
FKGRMKQQLLLLEDVGLGRSDLTITARPQVYVNYLIPKKKAVIAGAGTLRLQAKLKEQ
RLIQAADAKDSERIAQALKDIVLEFQVRVDPDMNMGVSTIADIIAEAAKNTFLVRKN
FPHAYIAKNLGGKNIPLKKEEVTATLLVEVTSNEYVTVLAQKQTEENQEG

CpN_0954 1088259 1089708
ychB-Predicted Kinase
GRKVCYKDIMQYFSPAKLNLFLKIWKGRDFNFHELTLYQAIDFGDTLSLKNMSKSLSS
NMYELLSNLIWKSLEIFRRETQIHQVPSWHLNKSIPQLSGGGSSNAATAYALNEH
FQHPITITTLQWAREIGSDVPFFFLQEQH

CpN_0955 1088612 1089175
r19-L9 Ribosomal Protein
RATPNPYSNNIATLGSRRNRKRCFFSSGTLGKGRGEHLFSIKKLNHKKYVLYLDHQ
GEPEKAYQSLLPDYSTGNHACFYGENDEKSVFRITDLKNNKHMERMSPFESHV
LMSGGATLFCVYLEELEDQSKVSSQIHSLIKQTCQIPVSRLYREPHWYSLKQSTYKNSP
LBSFQPI

CpN_0956 1089545 1090909
r19-L9 Ribosomal Protein
LWHEMILPPYSYSLKIGAAVLFFCSILHFTLPWLYTLQCSYEHKKLVFPECWKRYARL
SELEIRLSRVEIVFLLWAVLPFWFLVTEGYRISMAFYNSRNYGPAVFMVILILLESRP
IVVFAELVLSIAKLGKTSPPKSMXNLMIAAPLLSCLLKETGAMIIGATLLMRHFVPS
SRRAFATMGLLFSNISIGGLTSYSSRALFLIFPAKWEHSFFLSHFANKAVAILIST
TITYYIFIRKAEKFPDIPSDKPSVEKVPWMIICVNIIFVGSILRSRPLFMGALLFY
LGRQKFTIFVQDPINLSKVCYVGLFYAGLVFGDLQEWVNLNMQGLSDFGYMTVSYTSL
IFLDNALVNYLVHNLVATDCYHYLVVAGCMAAGGLTVLSNIPNIVGYLILRSAPFSSSTI
HMGWLFGLGALGPSIISLGVFWLLKNVPEFLYCYFFR

CpN_0957 1093812 1090963
ide(ptr)-Insulinase family/Protease III
KIAFRNCKMVKLLCPILICTSLSTICEQQFVPMVNPQCPLQVTPAAADQKIEKICSN
GLBLIISDPNLPTSGAALLVKTGMNADPEYPCMAHPTHECVFLGNEKYPEVSGPGL
SENNVGHNAPTYPNKTVPVFSVEHSFADGDLQVFLHINPKFRQEDLDREKYAVHQEFA
AHELSGRVRRHICQOLVAPQGHPCARFGCNGASTLTPVTTEKMAEWFKLHYSPENMAIA
YTSAPLSKAKKQFQSKIFSQIPRSKNYERQEPFLPSGDTSSKLQVLYINQAQPTSNLEIYW
HIYESSHPILGCKYKALAEVLNRESKNSVSLNKLQTLTDLDVEFFRSSLTGFEFYISY
ELTEGDKKHYSQVIDSTFQYLRVYIQEHGIPNPTLEISTINALNYCSSKSPFLDLCKQ
IVSLMNYEDLSTYPIHSLVYKYSSSEDEALLNLVSDPEQARFVLSSKNSHWEETATQHD
PIFDMTYVVKALDGVQYDGKVSQKPLPKPLNFIPKEVTLPGVHLLKKEQFPAPALS
YQDQKLTLYHCEDEHYTAPKLSQIRSRSPQISRSSPQFLVATELYCLAVNDQLREYYP
ATQAGLSFTSALGGDIDLVRSGVTTPVALLNSITSLPNLEISYETFLVYKQLELY
QGALLNCPVRSGDELASQVMKETYNTTKLSALEKLSFSEFQAFASNLPSNVHLEVMVL
GNLSEQQKDYLEMLQVFTASRSHATKPFYELQSEIIEIHDYPLTANGMLLLQDK
SSPSIQGVKCAEMLEFWLHHITFEELRTQOQLGYMVGARYEFASRPFGFLYIRSDAYSP
EELLAKTSLFLNKVSASPEKFGISQEKFANIRKAYINKILEPEHSLDMMSALFSLAFER
PVEFSTPDLKIAIETLTYEEFLKYCCFLSNELGTQTSVYIRGTQKTS

CpN_0958 1094803 1093793
plsB-Glycerol-3-P Acyltransferase
IYRAIYMQFSRYLYAFDNOYLPBELYQKFSVPHQNYIDAATKKAADQAEVLCLQWVKV
IIEEDLNKPFIFPPYHKKIRAPIDLFRSLIDFSLVIDDKNSRILNHLRLKEIEEYIARGD
NVLLANHQTECDQPLMYALGKTHPELMEINIFVAGDRVTSPLARFPFSGMCDLLCYIS
KRHATPPELREELKLNHNSQSMQIKLTLNEGKPIFYVAPAGGRDRKNAGRLYPSFSP
EIEVFRLLAKANSQTHFYPAFLKTDIIPPPKPIENAIQGEQRAIFAPVFFNGAELF
FDALCCKEELIHCDKAQRTLRAEKVFSIVKNLYEEL

CpN_0959 1096376 1094799
cylE-Axial Filament Protein
AQNYVITRTRVMENEILLNIESKEIRYAHILKNQOLFDTIERKKVPOLKGNITYRGVNTI
IRNIQCAFINDERENGFIIHIDILENSKKFECMFDMVDVADLPEEAGEAPLLSSEAPIE
RFLKLLFVIVQVVKEDIGKARALTSNIGIPGRYLVLLPNSPHRGVSRKIEDIPMRBQL
KRLIHFEMQMDILICRTACTTATNEALINEADLLTWKTLKCYFSTYEQMLLYGET
DILKPAVITCIDKNYKRLIIDYATYQKQKMLKLYCPDASIKIEYRDSIPMERFENIE
KEIKDATHKKIWLNSRYLFPDKTEAMTIDVNSGCTOLCSQVEETLVQINLEAEAEIA
WOLPILPVKXLVITIDIMKARKNRVRLERLKEHMYDAARCTILCMSFGLVEMTRQR
NKEIMLWTLTLPYCNKALIKTPESVYIEERDLKVINKEHSHCLLVVHPEIASYM
KQEHIDRIMINLAKQKAKLQINTDQSVHILHYQFFSLITGESIDL

CpN_0960 1096575 1097102
r19-L9 Ribosomal Protein
ILLGVSYSNPNQKALVGGGFGSMDCVNLKLYIFRLKLPDTERITYSISPEYIREKGE
EELLNSPIVEGGGLRIGDQWILSLSLKTLQGLQCPVCNNFFSHSVCLPDLQORVISHDE
VGSQVDFCRPLIRQELLESDECECSQOQPERKNILKFLDRKKHEGNSPFYEL

CpN_0961 1097106 1097207
r112-L32 Ribosomal Protein
FTALAGLAVHNNIMLAAITPPTKYNLRLTGLHNNVYVWYVNHNTVDALNT
KKFLSLNESGALSSTVFSLSYEGRIKALVKDIOQYITTVYDHLDFEELVEDRPIKNI
PIRCINAVDCIGVKGSLRQVIRAVRVVCKPKDIVPFLDVRVSGLSQTRKLSKIKIP
AGIETITPLKEVAITVSR

CpN_0962 1097301 1098275
plsX-FA/Phospholipid Synthesis Protein
ILSDPMFVQIGIDLMCGDHSPLVWQVLDVLSQSSSTIPFAFTLFASEEIRKQIQEEFI
SLDQKFPFKPIIAENFVAMEDSPLAAIRKKSMAALGLDYLQEDKLDADFISTGTGALV
TLARAKIPLFPAVSRPALLVUCVPMRGHAILDVGANISVVKPEEMVGFARMGLAYRQCLG
DSKIPITGLLNIISSEERKTEAHRQTFRMLRETTFGEAFLONIESGAVFDGAADIVVCTDF
TGNIFLTAEGVEFLQRLIGDKLEADIQRRLDYTFYPSVVCGLSKLVKCHGKACGSS
LPHGILGSINLAQARLCKRILSNLI

CpN_0963 1098374 1103224
pmp_21-Putative Outer Membrane Protein
TLPRFVAMVAKKTVRSYRSFSSHSVIVAILSAGIAFEAHLHSSLEDLGVFNKQFEHS
AHVEAQITSVLKGSDPVNPSQKSEKVLTYQVPLTQGGSSGSLDLADANFLEHFOHLEFEE
TTFVFGIDQKLWSDLDLDRNFQPTQEPDTSNAVSEKISSDTKENRKDLEDEPSKSKGLK
EVSSDLKSPKETAFAAISEDELEISNISARDPLQGLAFFYKNTSSQISSEKSSFGQIIF
SOGSANGSLGFENLKAPKSGAAVSDRDIFFENLVKLSFISCESLEDGGAAGVNIIVTH
CQDVTLDTCATGLDLALRLVKDFSRGAVFTARNHEVNNLAGGILSVGNSKAIIVK
NSAEKSGGAFACGSFVYSNNENTALWKENQALSGGATSSASDIDIQNCASAEFSGNS
LIALGEHIGLTDVFGGALAAQGLTLRRNAVVCVKNTSKTHGAILAGTVLDNETISE
VAFKQNTAGLGLSANDKVIANNFGEILFEQNEVRNHGGAICYGCRSNPKLEQKDSG
ENINIIIGNSGAITFLKNKASVLEMTQAEDYAGGALGHNVLDSNSGNIQFIGNIGSS
TFWIGEYVGGGAILSTDRVTISNNSGDVFKGNKQCLAQKYVAPQETAPVEDSASTNK
DEKSLNACSHGDHYPKTEVEEVPSSLLEEHVVSSTDIRGGGAILAQHIFTIDNTGNLR
FSGNLGGGEESSTVGDLAIVGGGALLSTNEVNVSNQNVVFSNMTNCGDSSGAILLAKK
VDISANHSVEFVNGSGKFGGAVCALNESVNTDNGSAVSFSKNTRLGAGVAAPOQSV
TICGNQNAIKFENFVFGSENRSGGAIANSVNIQDNAGDILFVNSNGSGGAILLAKK
GSLVASGNSPRLTITGNSGDLILFAKNSTQTAASLSEKDSFGGGAITYONLKIKNAGN
VSYFCNRAPSGAGVQADGTCVLEAFGGDILFEGNINFGSFAHILCGNDSKIVLSA
VQDKNIIFQDAITYEENTIRGLPDKDVSPLSAPSLFNSKPDQSSAQHHEGTIRFSRGS
KIPQIAAQIGTTLALSQNAELWLAGLKQETGSSVILSAGSILRIFDSQVDSAPLPTENK
EETLVASAGVIGNSSPTPKDKADVTPLVADIISITVDSLSFVEQDGLTLPPEIIPK
GTLKHSNAILDKIIDPTNVGYENHALLSSHKDILPLSLKTAEGMTGTPPTADASLSNIDK
VSLPSITPATYGTGVSSEKMGDRLVVGWQPTGYKLNEKQAGLVNLWSHYDKTLDRAL
LKQEIFAHTIAQRMELDFSTNVWVSGGLGVDEQCNIGFEDGFKHLLTGVALGLDTQVLE
DFLIGGCGSEFTGKTESQYKAKNDVKSVMGAAYAGILAGPWLKGAFFVGNINNDLTL
YGTLGISTGSGWIKGFIAGTSIDYRYIVNPRFISAIVTVVVFVEAEVYRIDLPEISEQ
GKEVRTFKTRFENVAIPFGFALEHAYSRSRAEVNSVQLAYVFDVYRKGLPDLTILKDA
AYSWKSYGVDPICKAWKARLSNNTEWNSVLSYLAFFNYEWREDLIAYDPNGGIRIIF

CpN_0964 1104812 1103301
No robust homolog present in Genbank/EMBL as of 11/7/98
QSILESIIKIFYLHNSKMHMSNPSILFSPAELIAKYNLIPKTSPIYPRTELIIILENA
CQTRLNVAQVLPSSLSFMSKKILNPGCCSGGLPVLNLAIFITISVLPILLPVL
IVAGRLFMPLPKKIVEDLSEPTTEENVEIQFIPALQALLFEDNKLRSFKIVPSVGL
KAPLNNPFLNRLVAISPOESQEAEMRKIPDLCSQLKVLKSLGLVTEPWKHLKYFEGKL
EHDNSDPKKTFFILIKLLIEALTGKSLPPTKTPSTKEMQALFIASSCKTCKPVEVIT
RSLNRLYSINAEQNDQLLIIVQVEFERELMSIQGDDAEYRFAAQHQERYTEAIEQVL
RNESAALKQWVINTMKFFHGKNLGLVTEHLQDTLGAALTIRQTTVDTHQGEDADLSAAL
FLNKYLSNGNQLVNSVFKSMQKADPETKALIREFADLILYASLRLPQTSAPTEVFSTLLM
DPNTEYENKACIAYLLVVLKIIEL

CpN_0965 1106769 1104925
lpxB-Lipid A Disaccharide Synthase
KPSFSKVLGMLNIPSGLVYLLYPLGLFALFFGSAFSIQWLLSKKRKEVYAPRSFWILSS
IGATLIVHGTIQSQFPVTVLHVINLITVLRNLNITSSRPIFRATLVLMLASVVFVTL
PLVYVMEWMAENPIFHLPLPAQLSWHLIGCLGAIIFSGRFLIOWFYIESNTKDFPLLP
WKIGLGLGLLALVYFIRIGDPINILCYCGGLFPSIANLRLFYKQRSTPYLDTHCFISAG
EASGDILOGKLIQSISLYPNIRFWGVGGPAMRQEGQLPILNMEEPQVSGFAEVLGSLFR
LYNRNKLILKTLIKKHATLIFIDFPDFHLLIKLKRKHGVRGKIHYVCPSIWAMPKR
KRIEQLHDMILLILPFEELFKNTSLETVYLGHPLVEEISDYKEQASWKEFLNSDRPI
VAAPFSGRRGDISRNLRIOQVAFNLSSLSQTHQVVSSSSAKYDEIIEEDLKAEGCQHSQ
IIPMFRYELMRSCDCLALACGTIVLEALANQPTTIVMCRLPFFDTFLAKYIFKILLPAY
SLPNIIMNSVIFPEFISGKDFHPEEITATLDELNLQHGSKQKEDCRKLCKVMPTTQQA
SEELFKRIFDTLPVAV

CpN_0966 1108055 1106748
pcnB-2-Poly A Polymerase
LLITIMVCENNILSGRGLLELLKKSNITLTPTIYSVSNHNIKLDKFSPLHSVIKTLRK
AGYIAYIVGCGIRDLLNTTPKDFDICTSAKPEEIKAFKNCILVGRFRLAHIFRSKQI
IEVTSFRSGSTDEBVLITKDNLWGTPEEDVLRDFTINGLFYDPEHEIIDYTGQVNDLR
NRYLRTIGDPTREFKQDPVRLRLKILSRSPFTVETQTALEQATQRELKSSQARVE
ELIKMLNSGAANKFFQLLIENHLEILFPYMDKAFRNLPALEQATYATLKALDDKILKE
AEYDRHQMLAIFLPLVNFVNRVYKHKHPLVLSLTVDFYIKNFLEQFADFSSTCSCKNF
ILTALILQMYRLTLPITPKKALFVFNKLLHHTFRLEALSLEIRSVIYKPLDKVYVAVI
RHQTLKCKKDSHSQK

CpN_0967 1108431 1109865
mrsA/pgm-Phosphotransferase
FTAYKFAFICACRSEKIRIGIDFRNNMQQSVVPLQTDQVGRANIEPMITVETTVLLK
AVARVLREGRKIKHVVVCKDTPGLGMYFNALIAGLNMGIDTILVLPITPGVAFITR
AYRADNIMDAHNPYRDKIKFISLEKFKIDVLEQPIETMVSEADPILPREDHAGVK
NKRVLIDAMGRVVEPVKVFTRKPTIKGLKVLVLAHGAATYKALIVFERIDAEVICYCE
PTGINTNEHCALITQVQKAVIFHQAILGLALVLDLPIIMVDEKHKIVRUMILICA
QDLKRRALGKHLNVATIMINIGVILKLEKALGLVATTPVQDRIHVLHMLHIEVTLXQD
SCHMIFLDYNTNDIYVNLQVLRIMITEMLDITAPIVKQPTPINNAVREKIPLET
ILPILRTLRVQDPAHILKILPVTENICPMVPSHKHIVQVLAFAADVDAELI
TQJRE

CpN_0968 1109889 1111721
glcB-Glucosamine Fructose-6-P Amidotransferase
FTAYKFAFICACRSEKIRIGIDFRNNMQQSVVPLQTDQVGRANIEPMITVETTVLLK
AVARVLREGRKIKHVVVCKDTPGLGMYFNALIAGLNMGIDTILVLPITPGVAFITR
AYRADNIMDAHNPYRDKIKFISLEKFKIDVLEQPIETMVSEADPILPREDHAGVK
NKRVLIDAMGRVVEPVKVFTRKPTIKGLKVLVLAHGAATYKALIVFERIDAEVICYCE
PTGINTNEHCALITQVQKAVIFHQAILGLALVLDLPIIMVDEKHKIVRUMILICA
QDLKRRALGKHLNVATIMINIGVILKLEKALGLVATTPVQDRIHVLHMLHIEVTLXQD
SCHMIFLDYNTNDIYVNLQVLRIMITEMLDITAPIVKQPTPINNAVREKIPLET
ILPILRTLRVQDPAHILKILPVTENICPMVPSHKHIVQVLAFAADVDAELI
TQJRE

CpN_0969 1109889 1111721
glcB-Glucosamine Fructose-6-P Amidotransferase
FTAYKFAFICACRSEKIRIGIDFRNNMQQSVVPLQTDQVGRANIEPMITVETTVLLK
AVARVLREGRKIKHVVVCKDTPGLGMYFNALIAGLNMGIDTILVLPITPGVAFITR
AYRADNIMDAHNPYRDKIKFISLEKFKIDVLEQPIETMVSEADPILPREDHAGVK
NKRVLIDAMGRVVEPVKVFTRKPTIKGLKVLVLAHGAATYKALIVFERIDAEVICYCE
PTGINTNEHCALITQVQKAVIFHQAILGLALVLDLPIIMVDEKHKIVRUMILICA
QDLKRRALGKHLNVATIMINIGVILKLEKALGLVATTPVQDRIHVLHMLHIEVTLXQD
SCHMIFLDYNTNDIYVNLQVLRIMITEMLDITAPIVKQPTPINNAVREKIPLET
ILPILRTLRVQDPAHILKILPVTENICPMVPSHKHIVQVLAFAADVDAELI
TQJRE

DRMCGIFGYNQDGVSVLVEGLAKLEYRGYDSAGLAADVVEQELFIRKTVGRVQELSNLF
QEREITPASVIGHTRWATHGVPTIEINAHFPHVDEGRSCAVVHNGIIEFKELRRELTAAQGI
3FASDTSDEIIVQLFSLYYQESQDLVFSQCTLAQSGSVACALIKHDPHTILCASQES
PLILUGLKEETFTASDSRAFFKYTRHSQALASGEFAIVSOGKEPEVYNLELKKIKHDKVQ
ITCSEDASDKSGYGYMLKEIYDQPEVLLEGLIKQHMDEEGHILSEFLSDVPKISFKEITI
VAGSSYHAGYLAKYIIESLVSTPVHIEVASEFRYRRPYIGKDTLILISQSGEATDTLA
ALKELRNRRIAYLLGICNPESALAGVDHCLFLEAGVEIGVATTKAFTSGLLLVFLGL
KLANVHCALTHAEOSFQOGLQSLPDLQCKLHSHESLHWAQPYSEYDKFLGRRLLMYF
VMEVNAKLEKEIAIDIAVITVPMKHGFAVLAIVTQDDUIFVEMQNMMEYF
VPHAVIATAPRAVITDAVITVPMKHGFAVLAIVTQDDUIFVEMQNMMEYF
PRNLAKSVTVE

CpN_0969 1111803 1112999
tyrP_1-Tyrosine Transport_1
VYVMSNKLGGSLIAGSAIGAGVLAVPVLTAAGGFFPATFLYIVSWLFSMASGLCLLEV
MTWMSKSNPVNMLMAESILGHVGKISICLVYLFYLLIAYFCBGGNLCRVFNQCN
LGISWIRHLGLPGLFMAILMGPIMAGTKVIDYCNRFPMFGLTVAFGIFCALGFLKIQPSFL
VRSSWLTTINAFVFFLAPFGQSIITPTLYYMDKKVGDVKAAILIGTLIPLVLYLWVEV
VLGAVSLPILISQAKIGGYTAVEALKQHRSWAFYIAGELPGFFALVSSFGVALGVMDFL
ADGLKWNKSHSPSIFFTLFIPLAWAVCYFIVLTLCLKYAGGGAAVIIGVFPPTLIWVK
GRYQKQHREKQLVPGGKALFLMFLLIIVNVSYIYHEL

CpN_0970 1113452 1114648
tyrP_2-Tyrosine Transport_2
VYVMSNKLGGSLIAGSAIGAGVLAVPVLTAAGGFFPATFLYIVSWLFSMASGLCLLEV
MTWMSKSNPVNMLMAESILGHVGKISICLVYLFYLLIAYFCBGGNLCRVFNQCN
LGISWIRHLGLPGLFMAILMGPIMAGTKVIDYCNRFPMFGLTVAFGIFCALGFLKIQPSFL
VRSSWLTTINAFVFFLAPFGQSIITPTLYYMDKKVGDVKAAILIGTLIPLVLYLWVEV
VLGAVSLPILISQAKIGGYTAVEALKQHRSWAFYIAGELPGFFALVSSFGVALGVMDFL
ADGLKWNKSHSPSIFFTLFIPLAWAVCYFIVLTLCLKYAGGGAAVIIGVFPPTLIWVK
GRYQKQHREKQLVPGGKALFLMFLLIIVNVSYIYHEL

CpN_0971 1114693 1115415
ycca-Transport Permease
EGSMGLVDRDYIQDSRVQGTASRVYGMWTAGLIVTSCVALGLYFSGLYRSLFSFWWVC
FATLGVFFINSKIQTLSVAVGGLFLYLLTBLGFMFGTLLPVYAAQYGGGVIAWAFGSA
ALVFLGAAVGAFTKSDLTISKIMTALIGLLTLVFAVVMFVSMPLIYLLICYGL
VIFVGLTAADAQAIRISSTIGDNNTLSYKLMFALMKYCNVIMVFWYLLQIFSSGSR
D

CpN_0972 1116377 1115430
ftsY-Cell Division Protein FtsY
RCIANSLLPFSYLVFLLQLTLLAMFQFPRNKLSQSLFKKNISLDLIEDAESLFYEADP
GTLEFEELCARLRRTKADASTIKDLITVLKLESEGLPQASQSSQTRPIVLSLLGTNG
SGKPTTAALKAHYKERSESVMLVATDTFRAAGMDQARLWANLGCQGFVSGQPGDAAAI
AFDQISATARGYSRVIDTSGRLVHGNLKLKLSIVSVCQKALEGAPHEIFMTVDSIT
GNNAIEQVVRVHDVPLSGLIFTKVDGSAGKGTLPQIAKRLKIPTKFICYGESLKDLENEF
DLDQFLNKLFPPEVEKI

CpN_0973 1116346 1117527
"sucC-Succinyl-CoA Synthetase, Beta"
BQSKGLFMHLHEYQAKDLLASVDYFIPYVVSSEEGELLITKSGLDSAVVKVQVHAG
GRKHGGVIVAKSSAGILQAVAKLGMHFTSNQADGFLPVEKVLISPLVAIQREYVAV
IMDRKHRCFVLMLSKAGMDIEEVAHSSPEQILITLPTLSYGHYYSYQLRQATKFMWEVE
VMHGGVQLIKKLAKCFYENDVSLLEINPLVLTLEGELLVDSKITTIDNALYRHPNLEVL
YDPSQENVRDLAKQIGLSYALSNGNIGIVNAGLAMSITLILKLGHGNAANFLDVGGG
ASQKQIEAVSLVLSDESIVKLFINTFGIMDCSVASGLVAVMTRDQVVPVIRLEGT
NVLEKKEIVQSGIPQCFVSSMEEGARRAVELSM

CpN_0974 1117523 1118422
"sucD-Succinyl-CoA Synthetase, Alpha"
VCRFRMYFHSLSKNTPIITQGITKAGSFHTEQCLAYTNFVGGVTGPKGGTLWLDLVP
YDSVLEAKQATGCRATMIFVPPYAAAEALIEEAEGTLEIVCITEGIPVRMDLEVARMD
NSTBOLIGNPCGIIKPKGECKIGIMPYIHLPGNIGVVSRTGLTYEAVWQTLQKIKQGS
ICVIGGDDPLNGTSFIDVLQALEEDPYTELINMIGETGSAEEAAWIQAHTCKPVVAF
IAGTAPKQKRMGHAGAIISGNSGDAKSIQVLRSGVTVVSPAHIGKTVDAVLRKAL

CpN_0975 1119038 1119637
No robust homolog present in Genebank/EMBL as of 11/7/98
GIEEQVALSIAIKILKILALIFPLVLLAWIRYQLHANFSCVVPFGSVNQAYKCS
EAKIEEMDLDDLETLEWSSRLCRDQMTFANRLEELIQELRVSETTELISLGGKRNLR
LLTLTHFPNPKRSRVESVGHEVVPVFDRLKREEEIIGDGPITRSNEELWALLDHGTARG
IKHTLWFSIFFKYLQTQELF

CpN_0976 1120079 1121185
No robust homolog present in Genebank/EMBL as of 11/7/98
ILMLVYCFDPSVPTSEPHRLMAALDRWFFLGGHARILTEGNHYRAQFNMSISTVEKI
LKLISYLLIPVLIALLIRCFILSRFKCNMKKSDLSARVPHDVQPFNDPNOERLN
IKWNRVYSGIDVLMVPVDYLSRQFPGKEIPEARECENYSDGQFSESKTSYLRAMLT
DIVGYILSLDETWTNVILKIRAMCITFESPGKEADPNYSPRVTHHYFDESUKALARHV
LGEGNMVNRLEALIRTEKPGKEGECITKQFLDYCKKHLVMSCPDFIESLVDEKIREP
RATPSILNSAVCDVIDRCKQEHLLKAIINEANRRLFGMKNSSTMRGNQVLFYTFISPPKL
PPAASSVVF

CpN_0977 1121329 1123402
No robust homolog present in Genebank/EMBL as of 11/7/98
LYINQFANILKSSFLMEVYSFSPVSTSFQHRVMAALDNWFFLGGRRLLKVVSLDCNSGQ
ACEEYVPISTTEKVLKILSYLLIPVLIALLIRYLLHSNFTAKVSKFWLKTLLQGLIDIK
SFLIPGCHVNTMDSATLFPKAIIRLECKRVDEYHRLHSQKVVYFIPAKQLDRLRLTHWL
PEKTRKTEYVVRHMLAHVMGYLTSGCKERLQVQVDSRSSTLSGAELKYLQYRFIDHPSQ
GEFORLLNENITKGGEDKEVQSDLDMAFCQWQVPSIVLQCPFTSEELVMSQKLD
LDCIYVPEDEDEQKFLNLTLLKAVLHHQFEGISVASMGVIFLICPDLSALQIPFLRNQK

CpN_0978 1122654 1123693
No robust homolog present in Genebank/EMBL as of 11/7/98
KYFMEVYGHFAVTRTGTQHRVMAALDAWFFLCKHRLKVVSLDCNSGWAYQELYSISTT
EKVLLGLYILVLPVLIALLIRCLILNENFIDVEKEPFLKIREGLDIESCKLPSSYNO
VVSFIDWIKKDKKKRPIDVDYHTLHSGDWVFPVQKIPKQPSYVWFQKTRKRVY
KNMLDIVLYLTPKXGEMLYQISKEVYQSATGLDPERVLYQVLTQDQELQGEVQRLNEE
RATPSIGKQKVLIVSDVFCQCVWPKFLEVITQSPAFIELVVEVSGKLNDFLCEKAN
TLDQILRNKILRAVVIHCHVWDIKFVAGLIIYTEAQLQIPQDS

CpN_0979 1123971 1125443
htrA-DO Serine Protease
GIDMITQLRSLWLAIVLGGSSLLALPLSGQAVVCKESRVSELPQDVLLEISGGFSKVATK
ATPAVYIESFPKQAVTHPSPGRRGPYENPFDFNDEFNRRFGLPSQREKQSKSEAVR
GTGFLVSPDGYVTNNHVVEDTGKIHVTLHDGQKYPAVYICLDPKTDLAIVIKISQNL
LSFGNSDHLKVGWDWIAIGNPFGLOATVTVGVISAKGRNLHIADFEFFIQTDAAINP
GGPGLINDQVIGVNTAIVSGSGGYIGIGFAIPSLMANRIIDQIRDGQVTRGFLVGT
LADIAAFLDPTDADPTDADPTDADPTDADPTDADPTDADPTDADPTDADPTDADPT
DADPTDADPTDADPTDADPTDADPTDADPTDADPTDADPTDADPTDADPTDADPTDADPT
TGKLIISVPEPSSVAASGIAPQQLILAVNRQKVSIEDLNRTLKDSNNENILLMVSGQD
VIRFIALKPEE

CpN_0980 1126988 1125504
"similarity to Saccharomyces cerevisiae hypotnetical 52.9KD
protein
FVMLNHAKKHAKPYVLIFFSTKDKLSYCDIIFNNSCKGPMNLSKHFDINSANFLEEF
FISFNSISADSDHLQDCENCAHFLVDHVNKIFDLEWETPGHPPIIYASYKSDPLSPTL
MLNHYHVAQPAQLSDGKWDPPFLREENGNIYARGASDNKQCFYTLKALQHYYESQGNF
PLNIINWIEEGEESGSLALFTWLEKKKALRADYLLIIVDGGFLSEKHPVYISGARGIVSM
KISLEEGNKKMHSGVLGGIAYNTRALSEILSLHHDNSIAIEGYDDOLAPSDSDRPO
LPKSDTLRECEENLFRPQGEYASYSPEESALRPTVEINGISGGYTGPGFKTVIPYRATA
VYSCRLVNPQDPAKAHQVHHLKQVPSLKFYSYIEILPGSGRWSSANLPIVKVQIEI
YSDLYNEECLDPLVMPATIPIGPLGAAQTSPIICGTSYSLDDIHAAEEHFSMDQLKKG
LSICQLDLKPKIKE

CpN_0981 1127019 1129952
Zinc Metalloprotease (insulinase family)
YTESMAGGATYRNFIIKSKDLPEIESKLEAEHKTGASIMMIVNDEENFVNICFRCT
PQTSNGVAHLEHMLVCGSENYVPRDPFSSMTRRSLNTFINAFTPGDFTCYPAASQIPED
FYNLSLVYIDAVFPHLLTKQSLQEAURYEFNSNHLCTGVVFNEMKMGMSGEARLSE
ALNAAIFPSVTYGVNSGGEPRIVTLSHEDVRAFHSQYSINRCLFYFYGNIKPSRHLDF
LEKKLLRQATKLEQAVSVPLQKRFKEPVNRNIIITPVHQEEDKVLFGISWLTCSILEQ
ELLALHLEELIILMGTDASPLKSRLLKSGFCQKTEMSTENDIREIMTLVCKGCSGPAQAQ
LEALFASLEEIIREGISENIVEGAVHOLELRKEITGYSLPYGLSLFRSGLLKQHGS
AEDGLRHSLSSELNLSKNSDYLAKLIRKYFLDNPHFARVILLPDTLVAKDNDKQDQL
LLSVSEKLTDENKEIKQNVRELTESQEQKEDLNGILPNLALDKVPTSKEGFLPIKEGLS
QEWLHHECTNDIVFIDVLDIPPLSGEELVLMGLVLMQLGCGGSRYSKHELEFLE
HTGGVDVYSDFSPHANKNSFLSPSVIRGKALSSKSEKLCGIVSDMLTSVDFTDIPRIE
LHMQNEALTNMSVNSPMSYAVSMACSGNSITGMSYLTGLPYVYKRLTETKNFDQNDI
EAVVTLQRLTYKCFSGKQIVISGSAHNYQOLKDNKFYGLLDYLIIVPEWENPNSINLV
TSRGLHIPARAAFNALAEIGDIADYDHPDAAALTVAEILDNVNLTIRKQEGVSGSA
AANLSRGSFYCYSDRDEIATYKTLKGVSEIASGNFTKEDIYEGALGVVQGLDMFVAP
GSRSAVYFRLKSGRIPLVRQAFRRSVLEVTEKHEICMVMDKYLETQVETTLISFAGEEM
LRNNVLTLDKDFPIVPAI

CpN_0982 1131215 1129962
yigN family
KKELASVMNLPVSLACLLSGCVFVLGVFVSSSLYARKKRAFLEIKQLEHENQLLTSK
NLRSRHOEQLIEDFSNRLALSSHKLKIDMKKEAQNYPGDTSKSFQISLPIQTLTTFQS
LETFTKHADRGRLKEQISQLLAVEKKLEHETHVLTDLIKHPSGRGMRGQIQLERILEL
AGMLKYCDYSQTSQAQAFRADIIIRLPQDRCLIIDAKAPISDSYFSVEEDKGDVLDK
IKHILKTLKSKSYWEKPHQSPYVILFLPGESLFDNADRLAPLEMTIGASSNVILSSPLT
LLALLKTLIAYMKQENLQKQIQEVSLLGKELHRLQVVFTHFQIKGNLQTVQSYNDMT
SSFQVRLPTLRKFEGLETSSSHQIEEPTPIESLTSFPHTCDDITNLVAVIESLEKQD

CpN_0983 1132045 1131206
pssa-Glycerol-Serine Phosphatidyltransferase
KNPLYCEQKLOIDMAGLDLEARGKRRVTPNATAFGLCCGLFIIFKSVLRTSSSVEL
FHRQLGSLILLISAMIAFDSDGAIARIMKASAFQAQFDSLDAVTFGIAPPIALIAKSLD
GIYGVNFSSLLITSIYSLCGVLRVLRNLSQKTVDSKVPYGLPIGPIPAASASVS
LALFLASDFFPDLPAQLRVGLLSFALLFIGGLMISPWKFGVKHFRNVSSFLVVTIGL
AALFFSGLVDFHVEVFFLVSWLTVLVGFFIFSIYIRKKS

CpN_0984 1132370 1135510
"nraA-Ribonucleoside Reductase, Large Chain"
GKVMVEVEEKHYTVIKRGMFVFPNQDRIFQALEAARFDRSLETSSPLPKDEESIAQI
THKVVEVLAKISEGVVTVVERIQDLVESQVYISGLQVARDYIVYDRQKAEGRNSSSI
EAIIRDDGSAKFNMKISAALEKAFRATLQINGMTTPATLSEINDLTRIVEDVLSHG
EAINLEELIQDIEVKQLMVAGYVDVAKNYILYREARADARANKQDQGEFVPEETTVV
QKEDGTTLYLLRTDLEKRFWACKRFPKTTDSQLLADAMFAMLYSGIKEDEVTACINAA
RANTREPDYFIAAELLTSSLYEETLCCSSODPNLSEIHKHKEIYINLGEYRLNPOL
KQYDLNDALEVLDSRDQFQSYMGVQVLYDRYFNHLEGRRLTAQIFWVRVSMGLALNEG
KBDNFALWITFYNTLSTFRYTPATPTLFSNGMRHSGLSYCLSTVKDDLSHIKYVLSNAL
LSKWAAGGINDMTDVRATGAVIKOTNCKSQGVIPFIKVNDAITAVNQGGKRGAMCVYL
BNHWDYEDNLEKRNKTGDERRTHDINTASWIPDLFVKLEKGMMLTFSPPDVGHLHE
AYGLEVEKLYVEYERKVESGEIRLYKKVEAEVLRKMLSMLYETCHPWITFKDPSNIRS
QDVGVVVRNCLCTEILLNCSSETAVCNLGSINLVEHINRDLDEKLEKETISIAIRIY
DNVIDLNFYTPKQANLTHRAVGLVGMQDVLNLSYASQAEVFSDECSSEIAY
YALISGLLAKERGTYASYSGSKWRDGYLPDITELLKETRGHNVLDVTSKDKWTPVR
DTIQKVMRNSQVMAIAPTATISNIIGVTQSIEMPKHLFVKNLSGFEPTIPNTYLKKL
KEGLWDAMBLDDLKYFDGSLLEIERINPHLKKLFLTAFEIPEWTIECTSRQKWDIMG
VSNLLYAEPODKKLSNMVLTAWKGLTYYLRQQAATVYKSFIDINKRIQPRVMMKN
KSASTSIVVERKTPVCSMEEGECSCQ

CpN_0985 1135432 1136571
"nrdB-Ribonucleoside Reductase, Small Chain"
ISVHKYCGRKKNNPRLFNRRRLRILSITEKRAKMEADILLGKLRVEYSKGLVNCNV
DVNOLYIEYKAWHEHYLNGCANMMLPTEVIMAPDIELWYDELSKEDPRVLLNLGFS
TAESILVGNINVLAFKHTNTPCARQYLLRQAFCAVHTHTFYLIETLLELDEGEVFNAYN
RAREIPAKQDFOMLTVDVLDPNFQVQVTECHYKFKLWYLIINDLFFYQGFVMLLS
FHRQNFYKIEQYQYIIRDTILNMTIDLINIKENFCAWTHLQELVIALICKAVE
LEIYAYKDLPRGILALRMTQVYVPLADPRLERIGLPIYERNPFWMTMDLNK
RKNPCTRTVTEQVAVNLSW

CpN_0986 1136712 1137693
yghH productase RNA Methylase
EILEMPQDLRPPILAKRRRTIQDQVAVRKHIIHQNPCTVTHOEPONTTACLLC
KXELWVVAQKQDQVQVAVIQLPILPZKILWPMINIGIQNRIYATITFFQYVVP
DQPTQVAVNPDPWPKMKHKKHPIQPFQVLTIRLQCAVFAIATDQTYLLIRIA
IATHLAPMPTTYIKMTDQYJEWIILWMPYQJEPYEPFKKAGI

CpN_1054 1207010 1209466
No robust homolog present in Genebank/EMBL as of 11/7/98
SRWHRFIMQVLLSPQLPPPHQHSVGSISSPCKSLRVLAITFLVQMLLLI3GALFTLIGI
PGLGAAISFGIGIGLSALGGVLMISGLLCLLVKREIPTVRPEEIPGVSLAPSEEPALQA
AQKTLAQPLKELDQDLDITQIEVFACLRLKDKSKYESRSLNDAKKELRVDFVVDLTSE
IFELRQIVAQEGWDLNLFINGGRSLMMAESESLLDPLHVSXRLQYLPBGDVRGEGLLKKA
KEIVAPIMGSHCEITHKVAFAFRNYSAMAEKAFKALGAEESVYRSLTQSYRDKFLSE
KATPFWHITWELDAAACAPFLDAAALVAVVWTEGAVGADINLIDWMT
TELHPHIMBETPHHLEPHTLKLHPLNLAATTEFKPKPKENICAFANAPPLP
TVRWYDQEFQKAGERLEKLHALYPEVSVSIRENKIQETLSNLEKAYEAIENYRCVRE
QEDYWKEEKRAEAFRERGNKILSPEELESSLEQFDHGLKNFSEKMLEBHLIKLQKEA
TAENVKILSDAESRLIEVFDVQMPGRLEIEIKTLRMABLPLPTKKAFAKASQVNS
CAEMLEKVKPYCKESLAVVTSKERLVSLDEDLRRAYTECQKRFQDGSGLSEVRACREQL
RERIQEFETQGLDLVEKELLCSVSRNLTETCCVSGVKKEAPGKKFYAQYDEIVRVV
QSRWMTMSERLREGVQACNMLKAGLSEEDVKLEEEYWLRYREERKNKEKRLVGTKIVAT
QQRVAAFESIEVPEIPEAPEKPSLLDKARSIFTREDHT

CpN_1055 1209583 1210521
No robust homolog present in Genebank/EMBL as of 11/7/98
CKYVYHHSYPPPPDHSGVAFCLSKFRVLAITFLVGLVFLISGALFTLIGISGLSAAIS
FGLGIGLSALGGVLMISGLLCLLVKREIPTVRPEEIPGVSLAPSEEPALQATQTLAQ
PKELDQDLDITQIEVFACLRLKDKSKYESRSLNDAKKELRVDFVVDLTSE
IQEGLWYLCRIQEMRDIGSTLFMSQVSLFKWAKELQVDFVWDMTEFVQLQIM
DQEGWYLCRIQEMRDIGSTLFMSQVSLFKWAKELQVDFVWDMTEFVQLQIM
RRICDTRKVAFTFDNRNAYGAKTAFKAFGALETQVYKSMTESYREAFCEYKTKILRDE
EKILRICYLELRR

CpN_1056 1210482 1211228
No robust homolog present in Genebank/EMBL as of 11/7/98
GEDIKMLSRVEEIEMLRVIELPLLPKQALEKAFVQVNSYKAKLTKEPCFPRESPAYI
TSEERLQSLDQTLERAYKEYQKRFQEPSRLESSEVSGREHLREQVKQFETQGLDLIKEEL
IFVSDVLFKRMVSLVSTVHVPFMEFYEFELHRLRLRAQWMAAETYSKVRKAFPEML
KETLEKAKAPREEEYWLCEERKSKEKRLILNKIEAAQQRVKDLEPPPIKETQKQKRKE
YSFFLRLS

CpN_1057 1211467 1213596
CT356 hypothetical protein
IIHFYFNFAMPELYTNKLTITEKSPYLLLYAHTPVMWYWGAEAFHIAIENKPVFLSI
GCKHSRWQVMQLQESYTNPEIAAMLEFVYVNVKDEELPYVAKYGLDQALMAVSDQH
ETVSWPLNVLFTPDLPVFFSVNYLGNBGLGLPSFPQIIDKLHFMWEDAEREALVEQAM
KVLIEASFLEGCVRKEIDESSLKRTVAALYQDIDPHYGGVKAFFKRLPGLLLQFFLYRS
LEFQESRGLFFVDRSLSMVALGGVHIGGGVSYTIDDKWLIPAFKRLIDNALMALNY
LEAWACLGKCEYRGIGKQILSYILSELYSPEVGAFFSYSSQAENMGAGGQNFYTWVSVEIS
NAGGEDAEIPCDVYGISREGFFNGRNLILHIVPHREIEELSEKHYRSIEATEDIVDRSDI
LKQIRAQSRHSRSDLSLTFNNGMIYTFAYAGRLGVEVEYIEIGKCGEFVNSLYKH
ELRRWRREGAKYRASLEDYDALIGVLALYSEGGCSFWLSFAEELMQEVVLSFRSEGG
FYSVDGRDSTLLIKQSPSLDGETISGNALICQCLLSLHLIEKKHYLTAFVLEILQIAQAC
PHTFSSGLLIIASONYFSRKHVKVLTALGDQEDRSPVLKCLGLFLPYLSLIWMTQEN
QEHLETVLPEYEHCLIPKGDCTATTIVVLEVDQCKRFKDELFRYRLISL

CpN_1058 1213742 1214836
CT355 hypothetical protein
EVNKLQYTLRGIVLVSTGCIFLMGHGYAAEVFVTSYENLLESKEQDPSGLAIHNRIL
FKVDEENNVLDVHKLNLFLYNSYPHLIDSPPARSQYAMVNVLESVIDEFLMVD
AKAKRIATDPTAVNQIEEMFGRDLSPLHYAFEMSNDIFNVDRTLTAQRVMGMVRSK
VRLKVTGKIREYRYKLEEEASRVYIKYRVLTIKANTESLASQIDKVRARLNEAKTWD
KDLTALVVISGGQLVCSSEFSRNSLSQSHQELDLIGYPKELCGLPKAHSKGYKLYM
LLDKTSGSIEPLDVMSKIKQHLFALAEASVEKQYKDLRLKRYGYDASMIKALLSEEA
PFLSSE

CpN_1059 1214848 1215678
kgsA-Dimethyladenosine Transferase
VTRSSPAQLSRFLSEIQNKPKKLSQNFVLDQNIIVKKIVATSEVLPQDWVLEIGPGFAL
TEELIAAGAQVIAIEKDPMFAPSLVEELPIRLEIDACYPLDQLEKYTLGKGRVAVNLP
YHITPLLLTKLLEAPDFWKTVTYVMQDEVARRIVAQPGGRDYGLSITFLQFADIVHVP
KVSACFCYKPKPQVQSAVIHMKVKETPLSDDEEIPVFFTLTRTAFOQRRKVLANTLKLGLP
KEQTEQALKELGILLNVRPEVLSNDYLALPHKMQAG

CpN_1060 1217694 1215727
dxs/tkt-Transketolase
YKRLPYIHITKVMYSSSCPLLDLILSPADLKLSISQPLGAEIEIRYRIISVLSQTGGHL
SSNLGIVELTIALHYVFPKSKDFIDFVGHQTYPHKLLTGRNNEGFDHNRNDGLSGFTN
PTESDHLDFSSGHAGTALSALGMAQTPLLESRTHVIPILGDAAFSCGLTLEALNNISTD
LSKPVVILNDNNMSISKNVGAMSRIFSRWLHHPHAKNLTQVEKWLAKIPRYGDSLAKHS
RRLSQCVKNLFCPTPLFEQGLAYVGPIDGHNVKLIPIIQSVNRNLPFPILVHVCTTKGK
GLDQAQNNPAKYHGVANFNKRESAKHLPAIKPKSPFDIFGQTLCELGEVSSRLHVVT
AMSIGSRLEGFQKQKPFERFDVGAIEGHAVTFPSAGIAKAGNPVICISYTFVHRALDNV
HDVCMQDLPLVIFADIRAGLAYGDRSHHGIYDMSFLRAMQMIICQPRSQVFLQQLYSS
LHWSSPSAIRYPNIPAPHGDPPLTGDPNPLRSQNAETLSQGEVDVLIAGLTCTFALSISIK
HQLLAYIGSATVVDPIFIKPFNDNLSLMLSHSKVITIEHSIRGLASEFNNEVATFN
FKVDILNFAIPDTFLSHGSKALTKSIGLDESSMTNRILTHFNRSKKQTVGDVRV

CpN_1061 1217932 1217666
CT330 hypothetical protein
FGSLMVEIHKKDPSLKKLFALQOSLETNLNSLDIVATYEAFLIYEGLNKALRKDQLCY
LLSVNSKGELKSPSGDPIVQTPPIHPPH

CpN_1062 1219835 1218159
xseA-Exodoxyribonuclease VII
KGFPMVMSPPQVACGLTERIKTLLENFNCQIIVKGELSNVSLQPCGHLYFGIKDSQAFNL
GAPFHPKUKYDYDRKPKDQDAVITGKLAVYAPRQGIIVAHALVYAGEGDLQKFEETKR
KLTAEYGFATEKKKPLPFAQCIGVITUTPTAVIQDILRLVLRARRNYKILVYPTVQGN
IAAHEIGKAIEVMAENLADVLIIARGGGSIEDLWAFNEELIKVKAHASTPIVSAVGE
TDYTLAFASDVKAPTFAAADIWCKSSSEGVQVFEQYLRHLHLHSGRQLLTSKKQQLLW
KHLDRAEFTYTAQQQLDSIEIAIQKQVQKTHESKQYDNITRWLQGLDQVSPMTCRLQS
IKKMLJQALSHKALQLQVRCHQLKGLSTYPRQIQQAQKLLHWRQDLDTLCPRLHYQKE
EYPHKHTPLKHAHNVLEQLRSLIVOKILGLGRLTPOCELNQKQKIAANVKETLATIL
HRYENGVARYZALKREQLHLNPNKVLKRGYAMLDFTHENAMIGVDSLQENARVRIDQLQ
IXEAILVTNITETCKLIK

CpN_1063 1219900 1220712

tpi3-Triosephosphate isomerase
FCRESMRKIFRENKRMTRQSYVLGNWKKHKTIQEAKYVQFLAQLLQEGFLCTIGIA
GPTFLRAIHVEVINTTGAFLLWLAGQNVHPELSGAFTTEIIPMLKEVTEVFVLVGHSE
HIFGESDAFIASKVKSVAQACLVPLVCGSELEVRREGKAHVIKKQLLLGLEQMDNGSE
FLIAYEPVWAIGTKVAEASDVQDTHMFCREVAERFSEATAEEISILYGGSVKVDNAQR
FQCCSDVQGLLVGGASLEQCSFPFEVAKNFV

CpN_1064 1220716 1220895
CpN_1064 1220716 1220895
CpN_1064 1220716 1220895

CpN_1065 1221140 1220928
No robust homolog present in Genebank/EMBL as of 11/7/98
RHRLGRHRRSTDPCFLFYFSIPEESLPDSCRLNQMPKHEHLSILLKKPIIDYKLTIS
YEKAIPTNGLP

CpN_1066 1221132 1221488
No robust homolog present in Genebank/EMBL as of 11/7/98
SMLNKIEIGMTVLFYAFILFIFLFCVLICGLILVQESKSMGLGSSFGVDSGDSVFGVSTP
DILKKTVCWCAVAFICIGLLSFTNLGLKGLAKEFLPLAAEESDTQASSESVDEADES

CpN_1067 1221675 1222292
def-Polypeptide Deformylase
IQVLVVRDFTTELCQAHVQTMIRRLLEYGSPILRRKSSPIAEITDEIRNLVSDMCDTMEA
HRGVGLAAPQGVNSLFLVMVCDRETDGELIFSESPPVFINPVLSDPSPTPIIGKEGCL
SIPGLRGVEFPQKITVTAMDNGKIFTEHLEGTARIIMHETDHLNGVLYIDLMEEPKD
PKKFKASLEKIKRRYNTLSKEELVS

CpN_1068 1223267 1223365
rnhB-Ribonuclease HII
MSCMPPPPVVTLTSAQNNLRDQLKEKNFISQPPQNTVFQARSNTVTCTLYPSGKLVIQ
KGSEEFIEFFLEPEILHTTHARVEQDLRLPRGLVDSESGKDFGFLPLCAAVASNAEILK
KLYENKQDSKNLKDTHIASLARIIRSLCVDVILYPEKYNELYGKFNLTLLAWAHA
TVINNLAPKPGDVFAISDQFAASEYTLKALQKKTETITLIQKPRAEQDVVVAASILA
RDAFVSIQKLEEQYQVQLPKGAGFNVKAAGREIAKQKGLLAKISKTHFTKTFDEICSG
K

CpN_1069 1223507 1223941
yfgA-HTH Transcriptional Regulator
VIMQEHIKHKLHLHGEIFRSSRESQSLSLKDVEAATSIRYSCLEAIEQCGCLKLISPVYA
OGFKIKYATYGLDGDSDILQEHYPMKIFKESDHNMEMLLDLESMDGRNSPERAHSWS
NLWAGLIIIGIMVWLGSLFSIF

CpN_1070 1225523 1224144
No robust homolog present in Genebank/EMBL as of 11/7/98
RRSLMTFPCGNCNYYRETPPPNPGGEDIPLEGQSGSGSGGRVITQPGTGGREMGISL
GSDNVLGMVEAGSLLNLLDSARMQRLGHYCYRTGTFWCREHCPGLQWIGGCCACCL
ETVDDPDNPELQQLIQYQVPCVMSFQGLPHCTQKIEBQGEPLDGDGKQEVENGCKL
HRELLKAAQPRCMGESLVLKLLQNMGLGEDMOQTPPWSLILQAVSSEGLSFTVSSDNPTC
WILQPEQQCPPPPTDEEQLGAGVGAAPQKQKHPAQECRVTKLNFRTLLQKLSRLV
LSLESYGKPLQGAQKIVDLIKKSLKRLVASDLATLPGIGLSLESQVFEVLVLLCLL
SKGLPLDPLHPEQTVLDPRVQGPQRILRKVLVTTTAGENIWRQTQGEAPRQAPFPDP
WDDDEIERDGIIVTGGGFGICQCLRCWRKLPTKRPNRWL

CpN_1071 1227336 1225885
No robust homolog present in Genebank/EMBL as of 11/7/98
KGMTFPCPNNSWFRMCGNFCEWVEVTTTEETTRQSASDISEAGSSGGAAPITTOPTKI
TKVEKRVQNTAQDESTIHMIEQAGELVDSILSHRRTQCTEYCYDSYATCGGRCGSF
GRLICGTYKACCLDREDNQVAGLVHECEQTHGPIAVALAAKTMGLNLMELVEKNTILSEE
QKNEFRQHCSEAKTQLYGTMQSLQNFLEGVNSIRERGLDLSLVOALVSIATRSWEKT
IESEASQTSASANSSTRIPACYILNTSPLTSLRSCGSRDARRPSSVGAEPQVYAKKNV
NGMAROLQKIQVNTLKTGDFSAFGPFLGLIVKMLNSFLLSASQSTSSILKHTGGEICYTC
PNFRDILVLLMAIGYCPANTDETSDVGIHMLDIPMTIFYRLQYSYRTGKTSASFGLK
PSLVRQESLDCPTAESVPLMSSLEEDENEDDDGDLAYQQRILECSGHLQTLFLGIK
INKE

CpN_1072 1227924 1228835
No robust homolog present in Genebank/EMBL as of 11/7/98
KKDYILHANWCCWQMLKIQKRMVSVVITVGAIVGFNSADAAPKKKKIPITQILYSPT
KVSSYLKNEDASTIFCVVDVDRGLLQHRVLSGPGWQETRRRQLFKSLENQSYGNERLGEET
LAIDIFRNKECLESEIPEQMEAILANSALVLGISSFGITGIPATLHSLLRQLNSFQKRS
IASSEFLKIDSPASDASVYKGVLFREGTATVDALSQFLAQDLSPKKIIFLEGDEPVV
QAVGSACIGWMNGLFLGVYPAQESLFSYVHPYSTATELQEAQGLQVISEVAQLTLNAL
PKMN

CpN_1073 1239011 1229832
Predicted OMP (CT371)
MRRYLFMVLLALCLYRAAPLEAVVITKIDQAQVLKFAREKTLVCFNIEDTVFPKMVGQS
AWLYNRELDKTTLSSEQAQAEFLWMMGISFLVDVDELVCANLRNVLTGLSLKRSKVLGI
SORPVLTKNTLRLIRSFNIDFTSCPAICEDGWLSHPTKDTTPOQAMAIENKILFVGSK
NQQPMDAALEVLLSGISSPPSIYVDQDAERLLDIPMTIFYRLQYSYRTGKTSASFGLK
SYNPKLTAIQWQIRKNSDEYYESLLSYVKS

.....
RNA SECTION
.....
rnpA 133193 133074
Ribosome Exit P RNA 607312 607649
16S rRNA 1000564 1002117
23S rRNA 1002415 1003778
5S rRNA 1003393 1003609
.....

tRNAs

tRNA #	Begin	End	Type	Codon
1	89657	89728	Thr	GGT
2	90998	91070	Trp	GCA
3	100000	100000	Met	ATG
4	100000	100000	Met	CAT
5	296075	296147	Val	TAC
6	296151	296224	Asp	GTC
7	409848	409922	Pro	TGG
8	462141	462214	Arg	CCT
9	672236	672318	Leu	CAA
10	677264	677337	Arg	TCG
11	739403	739486	Leu	CAG
12	781610	781680	Gly	TCC
13	784822	784896	Glu	TTC
14	784922	784994	Lys	TTT
15	836119	836191	Ala	GGC
16	843926	843999	Pro	GGG
17	877400	877473	Arg	ACG
18	1085605	1085676	Gln	TTG
19	1142034	1142118	Ser	TGA
20	1175863	1175944	Leu	TAG
21	1230028	1229942	Ser	CGA
22	1137462	1137389	Val	GAC
23	1030603	1030533	Cys	GCA
24	1000022	999949	His	GTG
25	961607	961536	Gly	CCC
26	807413	807341	Arg	TCT
27	786780	786708	Thr	CGT
28	715971	715889	Leu	TAA
29	708441	708354	Ser	GCT
30	680259	680178	Leu	GAG
31	631445	631373	Phe	GAA
32	626987	626901	Ser	GGA
33	293477	293405	Thr	TGT
34	293399	293317	Tyr	GTA
35	269142	269070	Ala	TGC
36	269065	268992	Ile	GAT
37	164389	164318	Asn	GTT
38	87522	87450	Met	CAT

What is Claimed is:

1. An isolated nucleic acid encoding a *C. pneumoniae* protein as set forth in Table 3.

5

2. The isolated nucleic acid of Claim 1, wherein said nucleic acid has a nucleotide sequence of an open reading frame in SEQ ID NO:1.

10 3. A probe comprising a hybridizing fragment of an isolated nucleic acid according to Claim 2.

5. An isolated nucleic acid that hybridizes under stringent conditions to the nucleic acid sequence of Claim 2.

15 6. An expression cassette comprising a transcriptional initiation region functional in an expression host, a nucleic acid having a sequence of the isolated nucleic acid according to Claim 1 under the transcriptional regulation of said transcriptional initiation region, and a transcriptional termination region functional in said expression host.

20

7. A cell comprising an expression cassette according to Claim 6 as part of an extrachromosomal element or integrated into the genome of a host cell as a result of introduction of said expression cassette into said host cell, and the cellular progeny of said host cell.

25

8. A method for producing a *C. pneumoniae* protein, said method comprising:

growing a cell according to Claim 7, whereby said *C. pneumoniae* protein is expressed; and

30

isolating said *C. pneumoniae* protein free of other proteins.

9. A purified polypeptide composition comprising at least 50 weight % of the protein present as a *C. pneumoniae* protein comprising an amino acid sequence of claim 1.

5 10. A monoclonal antibody binding specifically to the polypeptide of Claim 9.

11. A method for identifying a polypeptide comprising an amino acid sequence of claim 1, comprising the steps of: (a) obtaining a sample of a polypeptide; (b) comparing the amino acid sequence of the polypeptide to the amino acid sequence of claim 1; and (c) identifying the polypeptide as a polypeptide comprising an amino acid sequence of claim 1 if the amino acid sequence of the polypeptide is identical to the amino acid sequence of claim 1.

ABSTRACT OF THE DISCLOSURE

CHLAMYDIA PNEUMONIAE GENOME SEQUENCE

C. pneumoniae genome sequence and analysis of the encoded polypeptides
5 and RNAs are provided. The *C. pneumoniae* gene nucleic acid compositions find use in
identifying homologous or related proteins and the DNA sequences encoding such proteins;
in producing compositions that modulate the expression or function of the protein; and in
studying associated physiological pathways. In addition, modulation of the gene activity
in vivo is used for prophylactic and therapeutic purposes, such as identification of cell type
10 based on expression, and the like.

SF 1040660 v1

Contig463
Length: 273254..

1 ATTG TTCCTG TAAGAACT TCCAAAGCGC ATTTAATCAT TTTTAGTAAA
51 AAATAAAAAT ATACTTTTAA ATGTTGAGAA AATTTT TAGC TAAACTTTAT
101 AAAGGGTTGT TGGTGAAACC TTTGGGTAC TCCTCAGAAC GACTTTGTGA
151 TTCTATAGTA TTAAAAGGAT CTTGGAGTAT AACAAGTAAA GATCTTTGAG
201 GATAGCGTAG GGCCGTATTT TGAATAGCGT CCAATAAAGC GCGTTTGCAA
251 AACGCTTGAG TTTGGTTGTC CCAATAGAAA GTGCCTTCTT TAGGAAGAAT
301 CTCTTCTGGA GGCACCTCAT AGACCGAAGT AAAGAGAGGA AGAGCAACGA
351 TTGCTGCATG ACTTTCTATA GCTGCTTTAA GGCAGTTCTC GTACGCTAGT
401 AAAGCTTGGC GATAATATTC TTGCTGATTT GGTAACCTT CAGATTTAGG
451 GCCGCATACG TGGCCTAAAA AGGTCGGAAG AATACTTTTC TTTTCTGCAG
501 AGCTTAAATT TAGATTAAAC GTTTGATCTA GAGCTTCGTT TGGAAGTTTT
551 ACTACTCTCA CTTCGGTAGG GGAAAAGGGG TCTTCTCCTT TTGCGGGACC
601 CCCTTCGCGT TGCTTGCATG TATCCACAC GCTTTTATCT TTTAGGGTGG
651 AGTAAAGGAT AGTAGAGAGG TTCGTTGCAG TGTTGTCGAT CAGATTCGTT
701 GGGCCTACGG GATTAAAGAT GATCCCTGTG GATTGATTTT TTTCGATCAC
751 TCTAAGTCCA GTTAAGAAAG TAGGCTGAAA TGGTTGAGAC GCATCTGTTT
801 GTATCGCTAC CTTGAACTTA GGGTTCAGGT GATTATTGTA AAATTGCATC
851 TCGTTTGAGT AGCAGTCTAC GTTTTTTTCT TGCCACGCTT TTCCCAAAGG
901 CTTGAAGTTT TGCTCTAGAA CTTTCTGCCA GTTAGAAGAT ACCTTTGAGG
951 TCATTTGGTG GTAGACTAAG AAGGTTACAA CTGAGAAGAG GGCCGTGGTA
1001 ATGAGAAGAG CAAAAATAC AGGGTTCCCT AATACTATCG TTAAAGAGAT
1051 TCCAGCCACC AAAGCTCCTA AAGCTAAAGA AGCTAGGATT GCAAGAGTGG
1101 ATATTTTTCG TATGGTAAAC TGTTTTTTAG GAGCAATTTC TTTATCCCGA
1151 GGCACATAGG ATAGTACAGA AACTTGAGAG CTCTCAGTAC GTGAGGGTCC
1201 TGACATAACA TTTTTTTTGT AAAATACTTT CTATAATTTT AACATATTTG
1251 TGTTTATCGA TCCGAGAAAA TTGGAGAGTG AGAGCGCATG TCTTGCAATT

1301 TAGAATGATC GGGGACGACA TCTAGAGCTA TGTAGACATT GCGTGCCTAG
1351 TGGGAGCAAA TATAGCGAGA TATAAAGTAT AAGGGAATTG CTGTTAGGAA
1401 GATAAAGGAG CACAAAGGGT GGATACATAG CCCAATAGCT ATGGTGGTAG
1451 CAATCAGAGC TATCCAGACG AGTGCAATCG CAATAGTAAC GAAGAGGGCA
1501 AGCTTGAAAT TATAGCGAGG ACGAGTAGCT GGGGAAATA GAGAGGGAGC
1551 CGTTCCATCA AAACCGGGAG TAGCTGAAGA AGCCATAAAC TATTAAAAAT
1601 TAAGTTTTTT TCGGAGCATA AAGCATTTTA AAGTAGTGGG GTCTTTTTTG
1651 TCACGGAGAT GTCCTGGACT TCCCAAGCGT TTCTAACAAA GATACCTGCT
1701 TTTGAGAGGA GAACTTTTGA AACTCCTGCA AGGTCATCCT TCCTTGGCAC
1751 CAGTAGGTTT TTTGAGAAA TCGCGGAAAG ATTTTGGCGA AAGCTCTTAC
1801 AGTTGAAGGG CTTGTGAAGA TAATTTTTTT GTATTTAGAT AAAATATTTT
1851 TTTTAAGTTT TCGCGGCTTC ACTGTGTAGT GAGGGTAAGA GAAAAAGTA
1901 AATCGATTGT AAAGAAATTC TCTGATCACA GGTCTTGCGA GGGAGGAGTG
1951 GGGGTAGAGA ATGCGGGCTG AAGAGGGCAG TGCCTGTAGC AATGGGAAGA
2001 TGCCTTCAGC GATTTCTTGA GTTGCTACTA CGTACTTCAC TTGTCCAAGG
2051 AAAGAGAGAA GTCTTTCTTT GGTGGACTCT CCTATACAGA GGTAGGTCTT
2101 TGTTTTTAGA GTGGCCTTAG AAAGAAGAGA AGTCATTCTG GAAAGGAATA
2151 GGTGAGTGGA TGAGGGACTT GTGAGAAATCA CATGGGTGTC TTGTGGAAGG
2201 AATTGAAGAG CACGCTTATT TTGTGGAGTG CTTTTTGCAT AGGGGAAGAG
2251 AGTTAGAATA GGCAAATAAT GAGCTTGGTA TTTACGAGCG GTTTTTTGAT
2301 TCAATCCTAA GTAGAGGGTC ATGAGTCTTT TCGGGTAAAA GGAAGGCTGC
2351 CTAAGTTTTT GTACCTTCAA AGGGATATAT TGAAAATAAT TTTTCTTTTT
2401 CCCTTGTTTC TTCTTGATCA TGCCTTGATT GACATTTTTC ACTTTGAAGG
2451 CTAGGCTGGT TTTTTCTGGA CTTAGAGGTT CTCTCTATTA AGGCTTCGTC
2501 TTTAGAAGTC CTTGCTAAAA GTTTTTGAGA AATTTAAGAA ATTCGCAATA
2551 GTGGAAATAT TTACAAAGGT GGTGCGGTG GTTTCGTTGT TGCATAAGTT
2601 TTTAGAAAAT GCTTCGGGGA AAAAGGGACA AAGTTTAGCT TCGACAGCGT

2651 ATTTAGCAGC TCTTGACCAT CTCTTAAATG CGTTTCCTTC CATTTGGGGAG
 2701 AGAATCATTG ATGAGTTGAA GAGCCAGCGT TCCCATTAA AGATGATTGC
 2751 TTCTGAAAAC TATTCTTCAC TTTCAGTGCA GTTGGCTATG GGGAACTTGC
 2801 TCACAGATAA GTATTGTGAA GGAAGTCCCT TTAAGCGTTT CTATTCCTGT
 2851 TGTGAAAATG TAGATGCTAT TGAGTGGGAG TGTGTAGAGA CAGCGAAAGA
 2901 ACTTTTGTCT GCGGATTGCG CTTGTGTTCA GCCTCATTCT GGGGCTGATG
 2951 CTAATTTACT GGCAGTAATG GCCATTCTCA CGCACAAAGT CCAAGGCCCA
 3001 GCTGTCAGTA AGTTAGGTTA TAAAACTGTA AACGAATTAA CAGAAGAAGA
 3051 ATACACTCTA CTTAAGGCTG AAATGTCTTC TTGTGTTTGC TTAGGACCTT
 3101 CATTAAATTC TGGAGGCCAT TTGACCCATG GGAACGTACG TTAAATGTG
 3151 ATGTCTAAGC TTATGCGTTG CTTCCCCTAT GATGTCAATC CGGATACGGA
 3201 GTGTTTTGAT TATGCAGAGA TCTCCCGGTT AGCTAAGGAG TATAAACCTA
 3251 AGGTACTGAT CGCAGGATAT TCTTCCTATT CTCGAAGATT AAACTTTGCA
 3301 GTTTTAAAAC AGATTGCAGA GGATTGTGGA TCTGTCTTGT GGGTAGATAT
 3351 GGCGCATTTT GCAGGCCTAG TTGCTGGGGG AGTGTTTGTT GATGAAGAAA
 3401 ATCCTATTCC TTATGCAGAT ATAGTGACAA CAACAACGCA TAAGACATTA
 3451 CGCGGTCTC GCGGGGGATT AGTTTTGGCA ACTCGAGAGT ATGAAAGCAC
 3501 TCTCAATAAG GCGTGTCTT TGATGATGGG AGGTCCTCTA CCTCACGTGA
 3551 TAGCTGCTAA AACAGTGGCT TTGAAGGAAG CTCTCTCTGT GGATTTCAAG
 3601 AAATACGCTC ATCAGGTTGT AAATAATGCT CGTCGATTAG CAGAGAGATT
 3651 TTAAAGTCAT GGGCTACGTC TTTTGACGGG AGGAACAGAC AACCACATGA
 3701 TGGTGATTGA TTTAGGTTCT TTGGGCATTT CTGGAAAAAT TGCTGAAGAT
 3751 ATCTTGAGTT CCGTAGGAAT TGCTGTGAAT CGGAATTCAT TACCTTCAGA
 3801 TGCTATTGGT AAGTGGGACA CTCAGGTAT ACGTTTAGGA ACCCCTGCAC
 3851 TAACGACTTT GGGTATGGGT ATCGATGAAA TGGAAGAAGT TGCAGATATT
 3901 ATTGTGAAAG TATTGCGAAA TATTCGTTTA AGTTGCCATG TTGAAGGGAG
 3951 TTCTAAGAAA AATAAAGGGG AACTTCCTGA AGCCATAGCG CAGGAAGCTA

4001 GAGATCGTGT TCGCAACTTG TTGCTGCGTT TCCCGCTCTA CCCTGAAATT
 4051 GATTTAGAAG CTTTAGTTTA GTTAGGAGAG ACATTATTTT ATGGCAGACG
 4101 GGGAAAGTTCA TAAATTACGT GATATTATAG AAAAAGAGTT ATTGGAAGCG
 4151 CGCAGAGTAT TTTTCTCAGA GCCTGTAACA GAGAAAAGTG CTTCCGATGC
 4201 AATTAAAAAG CTTTGGTATT TGGAAATAAA AGATCCTGGA AAGCCTATAG
 4251 TTTTGTGAT CAATAGTCCT GGGGGATCTG TGGACGCAGG TTTTGCTGTT
 4301 TGGGATCAAA TTAAAATGTT AACCTCACCC GTCACTACTG TTGTGACAGG
 4351 GTTGGCAGCT TCTATGGGCT CGGTATTGAG TTTATGTGCA GCTCCTGGAA
 4401 GGAGATTTGC AACTCCTCAT TCTAGAATTA TGATTCATCA ACCTTCAATA
 4451 GGTGGACCGA TTACCGGTCA GGCAACCGAT TTAGACATTC ATGCGAGAGA
 4501 GATTTTAAAA ACAAAAGCTC GCATTATAGA TGTCTATGTA GAGGCGACAA
 4551 ATCAACCTCG AGATATCATA GAAAAGGCTA TCGATAGAGA TATGTGGATG
 4601 ACAGCCAACG AAGCTAAGGA TTTTGGTTTA TTGGATGGCA TTTTATTCTC
 4651 CTTCAACGAT CTCTAAATAT TTTATCTATT CTGGAGCAGG AAATCGTTTC
 4701 CTTCTTGGTG AAACACTTCC TGAGGTTGAA GATGTTCCGT TCTTATGCCA
 4751 AGAGACGAGG GTTGATGGTT TTTTATATTT AAAGCCCTCT TCTTGTGCTG
 4801 ATGCGCAACT CATTATTTTT AATTCCGATG GATCACGTCC AACGATGTGT
 4851 GGTAACGGCT TGCCTTGTGC GATTGCTCAC TTAGCTTCTC AGAAGGGAAA
 4901 ATCGGACATC TCTGTATCTA CGGATAGTGG TCTATATTCA GGATATTTTT
 4951 ATTCTTGGGA TCGTGTGCTT GTAGATATGA CTCTCGCAGA TTGGAGAGCT
 5001 TCTGTTTCATC GATTGGAGTC GCGTCCTGAT CCTCTTCCCA AAGAGGTCGT
 5051 TTGTATCCAT ACGGGAGTGC CTCATGCTGT CGTAATTCTT CCTGAGATTT
 5101 CTACTTTAGA TCTTCTATC TTAGGTCCTT TTCTTCGCTA TCATCAGACC
 5151 TTCTCTCCAG ATGGGGTGAA TGTCAATTTT GTTCAGATAC TGGGACATTG
 5201 CCAGTTGCGC GTTCGTAATT ACGAACGTGG AGTCGAAGSG GAAACTGCAG
 5251 CTTGTGGAAC AGGGGCTCTA GCTTCTGCTC TTGTTGTGTC AAACCTCTAT
 5301 GGATGGAAGG AGTCGATCCA AATCCATACT TGGGGTGGAG AGCTTATGAC

5351 TGTGAGTCAA AATAGGGGAC GGGTATATCT TCAGGGCTCT GTAAC TAGAG
5401 ATTTATAATT AGATGTGATT TTTGATTTTG TCATGCAAGG ATTTTAAAAT
5451 CTTGTTTAGG GATAGATCTT GCTCTCTAAC TGGGATTTTT CTATAATCGT
5501 AATTTATGAT GACGTATCCT GTACCACAAA ACCCACTTCT TTTAAGAATC
5551 CTTCTGCTTA TGGATGCATT CTCTAAGTCT GACGATGAGA GGGACTTTTA
5601 TTTAGATCGT GTTGAAGGGT TTATTCTCTA CATAGATTTA GATAAAGACC
5651 AAGAGGATCT AAATAAGATT TACCAAGAAT TAGAAGAGAA TGCCGAGCGG
5701 TATTGTTTGA TTCCGAAGTT GACGTTTTAT GAAGTAAAAA AAATCATGGA
5751 AACGTTTATC AATGAAAAGA TTTATGATAT CGATACCAA GAAAAGTTCC
5801 TTGAGATTTT GCAATCCAAG AATGCCCGTG AGCAGTTTTT AGAGTTTATT
5851 TATGATCACG AGGCAGAGTT AGAAAAGTGG CAGCAATTTT ATGTAGAGCG
5901 TTCTCGAATT CGAATTATAG AATGGCTTCG CAATAATAAG TTCCATTTTG
5951 TCTTTGAAGA AGATCTAGAT TTCACAAAGA ATGTTTTGGA ACAGTTGAAA
6001 ATACATTTGT TTGATGCCAA GGTGGGGAAA GAAATCACTC AAGCGCGTCA
6051 GTTGTTGTGCG AACAAAGCTA AGATTTACTA TTCCAATGAA GCATTAAACC
6101 CTCGTCCGAA ACGAGGCCGT CCTCCGAAGC AATCTGCTAA GGTAGAAAACA
6151 GAAACAACAA TTTGAGTGA TATTTATACA AAAGTCCCTC AGGCTGCTCG
6201 TCGTTTCCTT TTCTTACCCG AGATTACTTC ACCCTCTTCA ATTACTTTCT
6251 CAGAAAATT TGATACGGA GAAGAATTC TTGCTAACTT GCGCGGTTG
6301 ACTCGTGTTG AAGACCAGCT GAATCTTACC AATCTTTCAG AGAGGTTTGC
6351 TTCTCTTAAA GAGCTTTCGG CTAAGCTTGG TTACGACTCT CTTTCTACTG
6401 GAGATTTCTT TGGTGATGAT GATGAGAAAG TGGTCACTAA GACGAAGGGG
6451 AGCAAGCGAG GCCGCAAAAA ATCTTCTTAA TCTTCTATTT TGTGAAGTAG
6501 TTTATTTTGA GACGCTGTTC TTATTGCTTC TTTACATGAT CTTATTACAA
6551 ATCTTTCTTA TTTCTATTTA TTGTTTTGTT AAAATTTTAA CAATAGCTAT
6601 TTATTATTAG TCATTTTTTT AATTAAAAAA CTGTTAAAAT TTTTAAAGCT
6651 AATTTAAGAA ACAGTGAATA GTTCATCATG TCATCACTAC TGAGCTGCGG

6701 AAGAATAGAG CCGACTCGGG TTACCTGTAG CTAAAGACG TATCTTGAGG
6751 ATACGAGTCA GAATCAGTTG AGCACACGTC TAGTTCGGGC AAGTGTATC
6801 TTTTATGCG CATTGTTGAT CATTTTGGTT TGTGTGCCCC TCTCTAGTTT
6851 GATTCCAAGC ATTATGGCCT TGGCGACCTC TTTTACGGTA ATGGGGTTAA
6901 TTCTTTTGT GATGTCACTT CTTGGTGACG TTGCAATTAT AAGTTATCTT
6951 ACTTATAGCA CTGTTACGAG TTACCGGCAA AATAAGAGAG CTTTGTAGAT
7001 TCACAAGCCC GCTCGCTCCG TTTACTACGA GGGGGTCCGC CATTGGGATT
7051 TAGGACGATC ATCTTTAGGC ACAGGCGAGA TTCCTATAGT AAGGACGTTA
7101 TTCTCTCCAT TTCAGAACCA TGGTCTTAAC CATGCCTTAG CTGCTAAAT
7151 TTTCTATTT ATGGAGCATT TCAGCCCTGA GCCACCGAAC GAGCCTTTGG
7201 TGGATTGGGC CTGTTTGATT CGGGATTTTA GGCCTCACGT CAGTTCTTTG
7251 TGCTTTGTTA TTGAAAAACA AGGGTCATCG CTGAGGACTA AGGAAGGCAA
7301 TACGATTTGT GAGGCTTTCC GCTCTGATTA CGACGCCCAT TTTGCTATGG
7351 TAGATTGCTA CCGGTTGATC CACTCTAAGT TGATTATAGA GAAAATGGGA
7401 TTGAAGAATA TCGATATCAT TCCGAGTGTC ATGGTTCGTG AAGATTATCC
7451 TAGCCGTCCT GGGGAGGGCT ATCGCGAAGG CCTATTACGT ATGTATGGTG
7501 GCAAGGGGGC TCTGTGACTT CCCTACTTTA GTTCCTAATG AGCGCTTGCC
7551 CATAGGGCCT TTCTTTGTCC CGCAGCACAC TTCCGGTGCG AAGGGTAAGG
7601 AGTTTGCTAA AAGGAATTTT TCTATAATTT CGGGATTGGA TGACATATTA
7651 AAATTATGTA TTCTTCAAAG GCGTCCTTTT GCTTTGCAGT GGGATAACCT
7701 CTCTGTGAAA AGTGATTATG AGGAGGCTGG GCCCGCTATT GGGATACGTT
7751 CTCTTGAGCC ACAAGTTTCT CAAATTTCTC CAGCCCACGG CCGGCTATGT
7801 AGTACTTTGG TCCAGTGGGC CCCTATCCTT GGTTCCTGAGG AGCAGCTAGT
7851 TTGGTTAGAA GAAACAATGA AGCGCCTAAA GTTTCCTAAA AGTTTAGGTA
7901 GTAAGGACGC TGTTATTGTG GATTCGGAAA TGGTTCCTGT GAACGCCAAT
7951 CCTACTCAAG AGATACCTGC AGCTTCCGAG ACTGTAGAGT CTTACCTGT
8001 AGCTCCAGGG AATACAACAG ATACCATGCC TGCAGCTTCG GGAACCTACAG

8051 ACACCACATC TGGGGTTTCA GAGGCTGCGG CGGCTGAGGC TGCCGTGGAT
 8101 TCTACACCAG GGACAGAGGA GGAGCCGAGT TTTTCTCTGA GGTATGCGCT
 8151 TGTAGTTCAA AATGTTCCCT ATCCAGAGCC GCCTAAAGAA CCTGAGGTGA
 8201 TGTTTACAGA TGAAGAAAAA AGTCTGATTT TAGAAGCTAC TCGTGCGCGT
 8251 CGTATGGAGT TGGACTTGTA TAATGGCTAT TTAGCTGATT ATGAACTTTC
 8301 TAAGGATGAA ATACAGAAAC ACGTTCCTGA TTTACCTGAG AATTGGCGTA
 8351 CGAATTGGCG TTGGTCGGAG AGGCTCTATA AATTTTCTT TAAAACAAAG
 8401 AAAGAAGGAT TAGAAGAAAT TTTCTTAAAC AAAGAGTTAG GGAATATGAT
 8451 TCTTGCCCGA GGGCTGGCGG CAACTCAGTC ACAAGCACGT ATTAAAGTAT
 8501 TCAATTCTTT AGTGGCATGG CTCTTGCAAA GCTTTAACGT AGGGAGGAGC
 8551 TGTACAGCTA AACCTCTTCC TACGTCAAAA CTAGACCTCT TTAAATCGGA
 8601 ATTCGAGTCT AAGCCTAAAA ATAACATCTT AACGGAATTT TTGGTGGCCT
 8651 CTGATGAGGA GATTCTCTTT AAGGGGCTAC GGGTCCTAGA GCCTGGAATC
 8701 GAAGGTTGGT ATGACCATCC TGATCAAGCT GGAGAGATTC GGTGCGTACT
 8751 CGAGGGTCTG GTGCAGGCTG GACGTATTTT TGGATATTGG GAGAATCAGC
 8801 CGTTTGGGAG ATTTGTCCTT AGAGGAGTTG GTGAAAGACG TACCGAGCTT
 8851 GTAGAGCTTT TGGAGAGTTT AGTTGCTTCT GGTGAGATTA TGCAGTTCTT
 8901 TGAGTCTTCG GATGAAGAGG GTGCTTTTAT TATCGATAAC GAACCTAGCA
 8951 AGACTGCTAT GCTAAAACAG CGATTTAAGA GTTGTGTCAG GACGAAGCTT
 9001 GTCGGGAGTT TTGCTGATGA GAGTCTTCCC AGAGGTAGGT TTACCATTTT
 9051 AGTTTAGCGT GGGGTAGAGC ACTCCACGAA TCTTAGGGAG CTCCTTGCGA
 9101 CCAAGCTTGG AGATCCTCCA TGTTTTATTG TTTCTCTAGT AGCCAAATCG
 9151 TAGCCGCTCC TAGGAACAAT TTTTCTTTT TCGCAATATA AAATCCTGAT
 9201 TTAGAGAATA GGTCTTCAAG ATCGTGGTCC TTTGGAAGTT GCTGGATACT
 9251 TTTGCTGAGA TAGCTATAGG CGTCGGGATC TTTAGAAACA GACTTTCCAA
 9301 TCCAGGGGAC GACAGCACGC AAATAGAGCT TATGGGCACT ATAGGTAGGG
 9351 TGTGTTTTTT TTGGAGGTGT GAGCTCTAGA ATGCCAGTT TTCCAGAAGG

9401 CATAAGCACT CGGGAGATTT CTTGTAGGGC TTTATGTGGA TCCGAGAGGT
 9451 TCCTGAGGCC ATAGGCCATC GCTGCTAGGG GATAAGAATG ATTCTCCAAG
 9501 GGCAGTTGAT TAATATCGCT ATGAATAAAA GAGCAAGAGC CCTGGGGAAG
 9551 GTGTTGTTTT GCAATGTCGA GCATTGCTGA GGAAAAGTCG ACGAGAGTTA
 9601 CTGATGCTTG AGGGTGTGCG GCAATATAAC GCTTCGCGAC TTTTCCTGTT
 9651 CCTGCGCAGA GATCCAGGAG AGAGTATCCC GACCCTAGGA TCTGGATCAA
 9701 AGAGCGATTC CAGAAATGGT GCATTCCTAA AGAGAGTATT GTATTTGTGC
 9751 GATCATACTT ACTCGCTATG GAATCGAAGA TCTTTTACA GTCGGGCTTG
 9801 TTGGTAGAGG GTTCCATAAT ATTCCCGGAA TTTTCAAAG CTTTCGTAGT
 9851 GTTCTTCTCC TAGACGGTAC TGGCATAGGG CATAGTATTC TTGAAGAAGA
 9901 GAAGGGGGCA GACCTGTATG TTGATGAGCT TCTTTAAGGA CTTCTTCGGG
 9951 TGAAGATTCG AACTGTTGGA GGGCTTCTTC CATCGCAAGG TTGGGTAGGG
 10001 GATGTTCTTT CCAAGAGGTG CTGTGTAGAA GAAGAGCAAA TACAAAAGGT
 10051 AGCTTTGTAA GATCATACCA CCCCAGGCA AGGTCATAGG TTACAAATCC
 10101 AGGAAGTACA GGATGTTGTA GCGCTGCATC TCCGATTAGG AGGAGGCCAT
 10151 CATAATTTTC AGGGGTTTGT CTGAGTACTT TTGTAGTTAT GAATCTTAGG
 10201 ATATGAGGAG TTGGGATGCG CCAGAGATGA CGACAAAGCA CTTTTAAGAG
 10251 TCCTATAGAG GAGCGACTTT CTAAAGTTGC GGCAATCCGA GGTTGCGGTG
 10301 AGTTAAAGAA AGTGGGAGCT GCATAGAGGT TTACTGAG GATACGTTGG
 10351 TTTGCTGCAA TTCCAAAGCC GGGGACATAC CCCAAGTTAT GAGAGATAGC
 10401 TCCTAGGGAT GAGGTCAAAG CAACATCGAG TTTCCCTTCG ATTAGCAAGT
 10451 TGAGGAGGTC TGCAGGGGGA GCAAGAACAC AGCGAATATC GTTCTTTTTT
 10501 ATGAGTTGTA GGGACAGCGG AAAGGAATTA ATATAACTTA CGCAGCCTAA
 10551 GCTTATACAT GGCTGGAGTT GGTTAGACAT GGCGTTCTCC CTTGTTGTGT
 10601 GATGAGGGCC GCCATTCCCT CAGCGTCCAT TTAAATAGGT TCTTTAGATG
 10651 AGGCCATCTG GAAAACCTTT TCCCCATAT GTGTTGAAGA AAGGTCATTA
 10701 GCACCACAGG AAAGGAGGTC TAGAGCTGCC TCAATACCTA GGTAATTCCA

10751 TAAGGCTTTC ATATTGAAA AGTTGTCTAA GAAGATTCGG GCTACTGCCA
10801 TTAAAGATTT TAGAGGGATG GCATGACCCT GGCCTGATTT TCTTAATCTT
10851 TTTCCTAGGA CATTATTTTC TTGGGCGAAT TTTAGAAGTA TGAAGTTTTT
10901 AAAGCCCTGA GTTTCGTCTT GTAAGTCGCG GACTTTTACC ATGTGGGTGA
10951 CGAGGTCTTC AGGTCCTTCT TTATGATAGC AGAGCATGGT TATATTGCTA
11001 TGGATTCCCA GTTGATGAGC CATCTTATGG ATGTTGAGAA AATCAGAAGA
11051 AGAAAGGCGT TTGGGAGCTA AGAAATTACG TATTTTGTCTG ACGAGGATTT
11101 CAGCTCCTCC TCCGGGGATG GAATCAAGAC CCGCATCTTT TAATGTGAGA
11151 AGAACATCGC GAATAGAAAG GTTATCAAGA TCTGAGAGAT AGGCATATTC
11201 AATGGCAGTA AGAGCTTTGA TATGGATCTG AGGATCGTAC TCTTTGATTT
11251 TAGTAAATAG ATCGGAATAG TATTGCAGAT TGCAGGAGGG GAAACAGCCT
11301 CCCACGATAT GTACTTCTGT AATTGGAGTT TTTATATTTT GGATTTGCTG
11351 TAGAAGATCA TCTGGGGAGT AGAGCCATCC TTTAGGGTCT CCAGGTTTTG
11401 CATAGAAAGA GCAAAATTTG CAGCTGAAGT CACAGAAATT TGTAGGATAG
11451 AGGTACAAGG TTGAGGAGTA GTATACAGTG TCGCCAACCC GTTGTTTGCG
11501 AACTTGGTCT GCAAAATTCC AGAGTGTGCG TTGATCTTCT TTATTCGTGA
11551 GGAGGAGGAG ATGAAGAGCG TCTTCACTGC TTAATCGTTC TTGGGCATCC
11601 AGTTTTTCGA ATATGGAGTA GAGGGGGGAA GTTTTAGGGG GCTGTGGGAG
11651 GCACGTCGTC ATTTGATGAA CACTTTGATG TACTATTCTC TCGAGATTTT
11701 GTAGCACAGT GCTCTGTTTT GTCACATGTT TTTTTTTGGC AGCAATCTGG
11751 TTTTTGACAC CCTTTAGAGA GGGGCCTGCC AAGCAAATGG GAACCTACAA
11801 GTAGAATACC CATCCCTAGA CCTAACAATA CTGTGGCACA GCAGATTACG
11851 AGAAAAAGTG TCATCATAAG AAATCCTTAG ATAGGATAGT TTCTTAATTT
11901 AAATCCACCC AGATTGGGGA ACTCCAGGCC ATAGCATTGT CTGCCTGAGT
11951 GACCCTGAGA TAGTAGAATA CAAAAGGTGC TTTACCGTTT GGATCTTTTA
12001 GGGTCACTGA ACTTAGGGGT ACCATATCAT CGTATTCATA GTCCAGGTTA
12051 TTGCTATCGG GGAAGAAGGT ATGGAGAACT TCGCCATTGC GGATGATTTT

12101 TACAGTCTTG AGTAGGGCAG TGCCTGCCAC ATGACCAGAG ATGTGACGGT
 12151 TGACGTTGAG TCCAGGTTTC GACCCTGTGG AGAGTTCGGA GCCCATAGGG
 12201 GCTGAAGTGA TGTTGAAGCT TAAGACGATC CTAGGTCTTG TTGTAGCGTA
 12251 GCAATGACGT GCGAATAAAG CTTCAACAAG AGACTCTCGG GTATATTTAT
 12301 TACAAATGAT AGCCGTCAAC CCTGGGGAAT ATTGCACTTG CGGAGAGTCA
 12351 AAGTAGTCTT TATAAATTCC TCGATCGTCG AGACCCCCAG CAACAAATCC
 12401 GAAGCGGAGA TTCTTCTTTA ATCCTTCAAT TACTGTACCT CGAGGATCTT
 12451 CGCTATCTTT ACCTTGGATA GGGAAGGGGT TGTTTAGAGC GGCTGTGGTT
 12501 TCTGAAGATC CCCAGGCATT ATAAATTTCT ACAACTCTTT CGAACTCGGG
 12551 GTAGAAATTC TCAAAGTCAA AACCATGTTC TTTAGAAGCT GTGAACGAAG
 12601 GAATAGAAAT CATGTCGTGG TTGACAGTGC TTTTATAGAG CTTGGCGAGG
 12651 GGAATATGTT TGTATTCTTT GTGTTTCGAG TGGGACTTTG TTTCCTTGGT
 12701 ATGAAGGATG TGACGCACTC CCTCGAGATG AGGTTCTCCG CTATATTGGA
 12751 ATCCGGATAG TGTGATGAAG CGATCTTCTT CATTAAAGTC GGAGACAGTT
 12801 TGATTGATGA GCTTCCAAAT ATCTGGAGAG AGGTTCTCTT GATTTTCGAA
 12851 TGATGAAGAA GCATAGAAAT TCAGAGCGCG GTCATCTCGG AAATAACGCA
 12901 TACAAGTTTC AATATTTTCT TCAGAGTCGA CGCGTTCGGA TTCGCCGTGG
 12951 AGGAGACCCC ACATAAGATT CGGGGCGGAG TCAGCGAAAC ATTTGATAGG
 13001 GGCAGAGATG AAAATTTCTT GTGTAGAGAG GTTTTTC AAT TGGATGCGAT
 13051 AAATTCAGG CTCATTGAAA TAGAGATTAG GAAGAATAAC AAAGCCTGTT
 13101 TCTGGGATGA AGAGCTGCCA ATTTAAATTT TCTCTAAGAT GCTCGTAGGA
 13151 AAGCTCGATT CGGGTCTCTT CAGGAGAGAA GTTGGTGAGG TTCCCGAATT
 13201 CGTCTTCAAA TCGCACGGTG ATATCGAAGC GTTTGTTTTT AACGACATAG
 13251 GAGGGAGTAA AGATCTCTAT TTTTTTTAGG ACGTTTCCGC GGATATCCAT
 13301 AGAGAAGACA TCGGGTTCAT CATAGTTTCC TTCTCCTGTA GGATCGATGT
 13351 AGAGGTAAAA GGGTTTGCGA CGTTGTGCGA AAAGTTGGGC TCCGTTCCCA
 13401 GCATCATCGA CTTGAGGATG GTTTGGAGAG GCTCCCATGA CAATAGTGAG

13451 GGTTCCTCCT ACTTGAAGTT CGTAGGGGAG AGTAACTCG AATTGTGGAA
 13501 CGGGATTGTC TTTTACAGGA ATGGCGGTG CTTCGATGAT TTCGCCTTCT
 13551 GGCATTTCTG CGTAGATTAC GTTTCTAGTT TGGGAGAGAT CTGTGCGGGG
 13601 GGCTTCCCAA TCTGTGGGTT TCCCACTTCC TGCTAAGTCA AATTTACATT
 13651 TGGTTCAGC TGGTAGTGGT GTGGCAAGGG AATAAAGAAA TTTCCAAGTA
 13701 GAAATTTGCC CTGCTCGAGC TATCGAAGGG TTAACGTAAC AAACAGATCG
 13751 TCGCATAGTA AGAGGGAGGC TTTATATGAC TTAAGCGC CATCATATAC
 13801 TAACAGTGAG GTTTTTCTCA ATCCCCGTCT TTGTTTAGTG TTTGTATCGC
 13851 TTCATCCACA GTATTGAATA TTTTAAAGTA AGAAAGGAAT CCTGTAACAT
 13901 AGAGAGTTTG TTCTATGGTT TTTGGGACTG TAGTCAGGAC AATTTTCCCA
 13951 GAATGTTGTC CTAATTGATG GTAGCTTGC AGTAGGACTC GGATACCTGC
 14001 ACTGGACATG TAATCGAGGT GAGCACAGTC GAGAATGATA TTTTGGATC
 14051 CAGCTGCTAG GGATTGGGAA ATATTTTCTT GTACTTCTGG AGAAGAAATT
 14101 CCATCAAGTT TTCCGTGGAG ATGAAAGATT GTTGTGAGC CGTGTCTTC
 14151 TTTTGGATA TCACTCATCT AGATAGTTCT CCTAACTATA CGGGAGCTTA
 14201 AGTTTTCACT CTGATAAATC TTTAGCTTTT TTGCAAAGAG ATTTTATTG
 14251 GTGATGTTTG AGAATTTCGA TTGGGGGGAG GCAGGATGGG ATCGTGGAGT
 14301 GCAAGGAAAT CGGTCCTAGG ATGTTTACAT TCTTAGGAGA TATTGAATTT
 14351 ACGTTTTCTT GGTGTGATTT TTAGTTTTCC GACATTCGT TCTGTGCAGG
 14401 TAATGATTTT TATATCGAAG TTTTCGTGAT GGATACGCAT TCCTTTTTGG
 14451 GGAACAGCAC CCACTTTATG GAAGACATGT CCTCCTAGTG TATCGTAGCT
 14501 ATTTTCATGA TCGATTTTCA AATTGAAGTA CTCTCAGCG TCGGAGATAT
 14551 TCATTCTTCC ATCTACAATC CAAGAGCTTC CGATTTTCTT ATAAGGAGTA
 14601 TTTTCTTGTA CGTCGTGCTC GTCTGCGATC TCTCCTATAA TTTCTTCGAT
 14651 AATATCTTCC ATGGTAGCGA TGCCTTCTGT GAATCCGTAT TCATTGACTA
 14701 TGATGGCTAG ATGGCGATGT TTTGTGCGA ACTCTTGAG AAGAGAGGAG
 14751 GCTTTTTTTA TTTCTGGGGC ATAGAATGGG GGTTCGCTAC TGAGGATATG

14801 GGTGCTGA GGTGCTGGCT GCTTGATAG AGCAGTAAGA GATCTTTAAC
14851 AAGAAGGATT CCTGTGATGT TGTCTAAGTT TTTTATATAA ACGGAACGC
14901 GACTGTAGCC TTCTTCGCTT ACGAGAACCA GAGCTTCTTG TAGTGTAGTT
14951 TCTTCGGGAA GTGCGAAAAT ATCTACTTTT GGGATCATGA CTTACGGAC
15001 AATGAGGTTA TCAAAAGCGG AGAGGGCTTC GGAGAGCTGG CTTTGAAATG
15051 ATGTTGAAGA TCGTACTTGT TGGTTAGGGC GGCGTCTGTA AAAGAGCAGT
15101 TGCAGTGGGA AGAGACCGAG TTGGAATACC GAAGCTAGAA AACGGAGGTG
15151 GGCGGTGGTT TCTTTAGGGA CTTTGTAGA GATCCATGGG GGGAGGAATC
15201 CGTAAGCTAT CAGGGCGCTT AGAGAGTATA GGGGCCAGAA TAGGAGATCT
15251 TTGTGAGCTG TTTTGGAGG GAGGAGGGTA TAGAGTTTGT TCCCGAGAGC
15301 TCCATAGAGG ATGCAGAGCA GCGTGGCGAG AATTGTAGGA GCACTGGGGA
15351 AGGGGGGATA CTCTCTTCCT TTATCTTTGA AGAAGCGTTG GTTTAGGGTT
15401 TTTAGGAATT TTGAGGATCC GTGACAGGAC GGTGCGTAA GCCCGAAGGC
15451 TAGGAATAGA AGAATACAGA ATATGGCTAA AAGAATATGG AGCATGTTAA
15501 GCTGTTAGCA AAGCATGTTT TTTCTTAAC ATACACAGGA TTTGATTTTC
15551 TTAACTCTC ATTTTCTCT TTTCTTCTGA TGAGGTGTCG TCGTATCCGA
15601 GCATATGGAG AATAGAGTGG ACGAGGTATC TCGAGATTTC TTCGTAGATA
15651 TCCTCTGGT TTGGGGATGT GTTCTCTAAA AACCTAAGAG CGGCCTGTGG
15701 GCTAATGAAT GCTTCTCCTA AACATGAGG ATAAGCGGGA TCTCCGGGAG
15751 CATCAATAGG CAGAGTGATC GTATCTGTTA GAGAAGGATC AGCAAATACC
15801 TTATCATGGA GTTCTGCAAG AGCTTTATCT TCTAGGAAGT AGATAAAAAT
15851 TTCATTAGTT GTTACTTTTA AGTGCTCTAA GAGCGTAAGA ACCAGCTTCT
15901 CTACAGAAAC CAAATGAATA GGAATACATG TTTGCTCATT GGAAACATGT
15951 ATTTTGATCT TTTCTTGCGT CACGCGAATG AAATTCCCAT AGAGAAAAAC
16001 AACTTATTT TAAAATAGGG GTCTTAGGTA AACCTGTGAC TTTTTCGCT
16051 GTACTATCAT TCCAACGGCC CAACTTACGC AAGACTTCTA CTCGCTCAAA
16101 ACGCTTCAAA ACATTTCTTT TGGTAACCCC TTTGACAGAT TTACCATAAC

16151 TACGATGTCG AGACATAATC CTGCTCTAGA AATAAACCTA TTTTCGGGAT
 16201 AAAAATGCTA TCACTGGTGC TCAATGCATT CGTATGCAAT TTATATACAA
 16251 TTCTTGGAGC TGGCGCTGGA TGCACAATGG CATACTTAGG TTTACGTTTT
 16301 TTAGGACTTT TCGCTCTGCG CCGAGCTTGT TTACTCATTC GTGTCATAAA
 16351 TAACTCGATA TTGAATTTTT TAATTTCTCC AGACAACGGA AATTGAGGAT
 16401 ACGGAGTACT TTATAAGAAA AAGGATAGTA AAAGAAGAGT TTTTTTTCAA
 16451 GAAGATGACG TCTTTTAGCT GCCTTGATCT TGGTGTAGCT CGTCAGGGAG
 16501 GGGAGACGAG GGGCATGCCA TGGGGGTGA GAGGACTCCA CAGGAGATAA
 16551 TATATTTGAA AGCATCTTCG ATTTTCATAT CTAGGAATAC GATATCAGAT
 16601 TTTCTAAATA GGGTAAGAAA CCCTGAGGTG GGGTTGGGTG TTGTTGGGAT
 16651 GAAGACCGTG ACGAGGGGGT CGTCTCCTT TTCTCCTGTG CAGCATACTG
 16701 TGGGTGCGTC TCCAGCGACG AGACCGATGC ATTGAACATT TCGGTTAGGG
 16751 AAAGGAACCA TAACTACTTG TTTGAAGGAT CCTGATTTTG ATCCAAATAT
 16801 GGTAGTCATG ACTTGTTCG CAGCTTTATA CACTGTTTTA ATGATGGGAA
 16851 TTCGGTGTA GATTTTGTG TAGATAGAGA GTAGGGATTT AAAAATCATA
 16901 ATTCTCGTGA GGAAACCTAG GAGCACTGTG GCGAAAAGA GACCGAAGAG
 16951 TAAAATGATT TGCAATACGA ATTTTAGAAG AGCTCTATGT TTAGTATAAA
 17001 AGCTAAATTT CTCAAAGAAT TCCGAAGCCA AGCCTACGAA GGGTTGGGTT
 17051 AGGAAGTTCA TGATCATAGT AACAAAGCA ATAGTAATTG CTAGAGGAAG
 17101 GAGAATAACA AGTCCTGTAA TAAAGTATTT TTTCATGATT CTCCTGCAAG
 17151 ATATGAGGAA ATGGGCATTT GTTTCTTTAC TATACAGCTT AAGATTATTT
 17201 AAGATAAAAC TTTTCCCGAA TCTTCTGGGG ATAGGAGAAA TCTCCATGGG
 17251 ACATCACGAT ACTCTTGAGC ATAATCGATG CCGATCCGGG CAGTTGCTGT
 17301 TAGAGTCCCA GAGATTTTTT CTTTGCTGAT ATAGAGAGCT GGGGTATTTA
 17351 GGCGTTGCCT ATTGTTTTCC AAAGAGATTC CTAGAGCTTG GCACACTTTT
 17401 CCGGGTCCAT TGGTGAGAAG GTGTGGGGGT TTATCTCTCC ATTGGCGGCG
 17451 TTGGATCATA AGTTCTTTGC CTTGATCAGG AAGGATGGCC CGGATCAGGA

17501 CGGCATGGGG AATGTCCTCA GGTCCAGTGA CAACATTCAA TAGGTGATGC
17551 ATGCCATAGC AACGGTAGAG GTAAGCAGAG CCTCCTTTCA GGTACATCGC
17601 TCTGTTCTC TGAGTTTTTC TGTAGTTGTA GGCGTGGCAT GCTTTGTCAT
17651 CAGGGCCACG ATACGCTTCG GTTCTACAA TGTAACCTGA AGTTATCAGA
17701 CCCTCATGTG TTGTGATGAG TTTATGTCCT AAAAGCTGTT GCGCTAGTGT
17751 AATTACATCT TCCGATAGAA AAAAATGTTC TTGTAGCACG TTACGAGGCT
17801 CTTTTTTTCG TTCCTTTTTT CTAGAAGGC GTTTTCTTTA TTTCTTAGG
17851 TTTATCTTCT GTGGTTGTCG CTATAGACCA GACGATTTTT TGCCTAAGGA
17901 GATTCACGGA ATCAATAGTG ACTTTTATAG AAGCTCCAGG TTCATTTTA
17951 TCTGGGATAG ATTCTGGAAG AGCGTTTTTC TTTAGGGAAT ATTCTTTAGG
18001 GAGTTCTGCT GCTGCAATGA ACCCTTCATG GCAGAATTCG GTCACTACAA
18051 ATGAGAGTCC TTCATGATTT GCAGTGATGA TATACGCATG GTATGTAGTT
18101 TTAGGTTGCT CTTGCAAAAA TTTATTTATG AACCGAGTTT TTTTGAGGTT
18151 TTCGAAAGAA TTTTCTGCTT TTGCGGATAC TCGTTCTTTT GTAGAGCATG
18201 CTCTTACGAT AATTTGAGG TCGTTTGGT CTATAGATAG GGGTTGAAG
18251 AGAAGCCTGT GAACAATAAG ATCGATATAT CTACGTATGG GACTCGTAAA
18301 GTGGGTGTAG TAGTCGAGCT TAAGTCCGTA ATGACCTTTA TTTCTGTAG
18351 AGTAGGAGGC TGTTTTCATA CTTCGGACAA ACTGCGAGTG TAGAACTTGC
18401 TCTAGGGGAT GTCCTGCTGA CGTAGTTTGC AAAAGGTATT GGTAAATCAGG
18451 TTCTTGTGTG GGAGTGAACG TGATATCAAA GCCCATGTTT TTTGCCAATT
18501 CTTGGAAGGC GAGTAGGTTT TCATCATTGG GAGGTTCGTG ACTACGAAAA
18551 GGTAGAGAAA CGCCTTGATG GGAGATATGA TAGGCGACCA CTCGTTTGC
18601 TTTAAGCATA AACTCTTCGA TGAGTTTATG GGAGAAGGTC TGGTGGTTTT
18651 CTATCAGAGC TACGGGTCT TGAAGATTAT CCAAGGACAT AGTGA CTGAG
18701 GGGAGGACAA AGCGAATGCA ACCACGTTCT TCACGGATAT CGGAAAACTT
18751 TTTACTTAGA GTGGCCATCT CATTGAGGAT TTTTGAGAGG GGGTGGGAGT
18801 GTTTCTTTTC AATGATGTTA TCGACTTCAT CGTAGGTCAT ACGATATTTG

18851 CTTCGAATGA CGCTACGGAA AATCTGGTAA TCTGAAAGAT GACCTGATTT
 18901 TGTAACGTC ATAAATACGG ATACAGCGAG TCTATCAACG TTTGGTTTTA
 18951 AGCTGCAGAG ATTATCAGAG AGTGCTGATG GCAACATGGG AATGACTTTC
 19001 CCTGGGAAAT ATGTAGAGTT ACAGCGTTTA GCAGCTTCTT TGTCTAGGTG
 19051 AGAATGTGGG GTAACGTAGT GGGAGACGTC TGCGATGTGT ACACCAAGAA
 19101 TGTAATTGTT ATTATGATCG TAGGTGAGGG AGATGGCATC GTCGAAGTCT
 19151 CTGGCTGTGG AAGAGTCTAT GGTGAAACAG AGGAGATCAC GGAGATCTTT
 19201 GCGAGAGTGG AGAACTTGGG TAATGTGTTT TTGAGAGAAA AGGCTTGCTT
 19251 CTTCAATGAC CTCTGGGGGG AATTCTTCGG CAAGGTATA TTCGGCTTGA
 19301 ATTGCCTGAA AGTCCGCTTT AGCGTTGGTG ATGTGGCCAA TAAATTCGAG
 19351 CATTTGTAAG GCTGGAGAGG CTCCTTCTTG GGGTTTATCT ACCCAGGGAG
 19401 GAGTGCTCAG AAGAATGCGA TCGCCGATTT TGTAAGTGCG TCCGGGAAGG
 19451 AGTTCTACTG GAATTAAAGA TTGGGATCCC GACATGCTTG TGTAGGCAAG
 19501 TGCTGATGTG GGA CTGACTA GTGAGGTGAT CGTTCCTACG AGTGTGTGTT
 19551 TTCTCTTGC GAGTACTTCG CTGATAGTGC CTTTGAGTTT TTGTCCGTCT
 19601 CTTGGATAGG GAAGCACGGA GACAATCACG TGGTCACCAT CTAGAGCCCC
 19651 GCGTAAATCT CGGGCGGGAA CAAAATATC AAATGGGTAT TCTTCGGGGT
 19701 TGTCGGGAGA AACAAAACCG AAACCTTTTC TAGCATGAAC AAATAGGGTT
 19751 CCTGGAATAA AAATCTTCAA GGATTTACCG TATGTTCTTC TCCCTGGTTT
 19801 TCTTTTGGT TTTTCAACA ATTGGGCTCC GCCTGTAAGT TTAGGACATT
 19851 GAACGAGAAA TCCCGTAGTT TCACTCGTGA ACTGAGTGGC TCAACAAAAT
 19901 TTTCCCTTTT AAGTGGGTTT TGTGATTATG AAAGGAGCAG TCTCAAATTC
 19951 AAAGCCAATT GACTAAAGA AGCTCTGTTA TTGCAACTCC TTGATCAGAG
 20001 AATAAGAGAA TGGAAGTGT TTCTGTTTAT AAGGAAGTTT CCCATCCTCT
 20051 TATGAGGATG GGAATAGAGA AACTTAAATT GAAAATTTG ATTACTTATC
 20101 GTCGTTATCA ATAATTTCTA CATCAGCTTC TTCGATATGG TCTTCTGAAG
 20151 AACCGTTATT TGAAGGAGGC TTCGTACTGA AACTATGTTT TTTCAAATCT

20201 TCTGTATTGA TGTTAGGTCC ACCTTTAGCA TTGGCTGCCG ATGATGCTGC
20251 TGCTGATGCA GACTGCGATT GCATAGACTC TCCAATTTTT TGCATATGCT
20301 TGCTTAGGTC TTCAGTAACC TCTTTAATTT TTTCAATAGG AGCGTCATCT
20351 TTGAGTGCGT TGC GCACGTT TTCGATTGCG TCTTCGATTT CTTTAACTAA
20401 AGTTTCAGGA ATTTGCTCCT TATAATCTTT AATAGCTTTT TCGGCTCTGA
20451 AGATCATGCT ATCGGCTTCA TTTT TAGCAT CTGAAGCTTC ACGACGTTTT
20501 TTATCTTCTT CCTTATTAAT TTCGGCATCT CGAACCATTC TTTGGATTTC
20551 ATCTTCTTGA AGTCCTGAGC TTGCTTCGAT ACGAATTTTC TGTCTTTAC
20601 CGCTGGCAAC ATCTTTAGCT GAGACATGGA AAATCCGTT TGCATCGATA
20651 TCGAAGGAGA CTTCGATTG AGGATGGCCT CGAGGAGCCG GAGGGATATC
20701 TGTAAGATCG AATCTTCCGA TTTCTTGTT ATCTTTGGCC ATGGGACGCT
20751 CTCCTTGAG AACTACGATG GTAACCGCAG CTGGTTATCA GCAGCTGTGG
20801 AGAAGATTTG TTTTTCTGT GTAGGGATTG TAGTATTTCT CTCTACCAGA
20851 GTCGTCATGA CGCCTCCTAG AGTTTCGATA CCCAGAGATA GGGGATAAC
20901 GTCTAGAAGT AGAACATCCT TAACTTCTCC GCCAAGAACA CCACCTTGAA
20951 TTGCGGCTCC AATAGCAACA ACTTCGTCGG GGTGACTCC TTTATTAGGC
21001 TCTTTGCCGA AGAGTTCTTT TACAGTTTCT TGC ACTGCGG GCATTCTTGA
21051 CATACCTCCA ACTAAGAGAA CATCATCGAT ATCCTTAGCG GAAAGTTTTG
21101 CGTCACTGAG TGCTTTGATG CATGGAGATT TTGTTCTTTC GATTAGAGAG
21151 GCTGCGAGTT TCTCGAATTG CGCACGTGTG AGTGTCAATG CAAGGTGTTT
21201 AGGTCCTTGT GCATCCATTG TGATGAATGG CTGATTGATT TCTGTGGAAG
21251 AGACTCCTGA AAGTTCTATT TTTGCTTTCT CAGCAGCATC TTTAAGTCTT
21301 TGTAAGGCCA TATTATCTTT GCTAAGATCA ATGCCTTCTT GTTTTTTGAA
21351 TTCTTCGATC ATCCATTTGA TAATGACTTC ATCAAAGTCG TCTCCACCGA
21401 GGAGAGTATC TCCATTTGTA GATAGAACTT CGAAGACGCC ATCACCGATT
21451 TCTAGGATGG AGATATCAAA AGTTCCTCCA CCAAGGTCGA AGACAGCGAT
21501 TTTTTTATCA CCGACTTTAT CGATTCCGTA GGCAAGAGCT GCTGCGGTAG

21551 GTTCTGGAAT GATACGTTTT ACATCTAGAC CTGCAATGCG TCCAGCATCT
 21601 TTTGTGGATG CTCGTTGAGA ATCATTGAAG TATGCGGGGA CGGTGATCAC
 21651 TGCTTCTGTG ACAGTTTCGC CTAGATAAGC ATCAGCTGTC TCTTTCATTT
 21701 TCATTAAGAT TTGTGCGCCA ATTTCTTCTG GAGTGTATTG TTTGCCATCA
 21751 ACTTCGAAAA CGGCATCACC TTTAGATCCG GAGGTGACTG TATAAGGAAC
 21801 GGT TTGGATT TCCGAAGCTA CTTCAGAGTA CTTACGGCCA ATAAAGCGTT
 21851 TTGTAGAGCC GAGAGTTTTT TCTGGATTG TCACTGCTTG ACGTTTTGCT
 21901 GGAATCCCCA CTAATTTCTC ATTACCTTTG AAGGCAACGA TCGATGGCGT
 21951 GGT TCTTGTT CCTTCGGATG ATGTAATTAC TTTAGCTTGT CCTCCTTCCA
 22001 TAACAGATAC GCAGGAGTTT GTTGTGCCTA AGTCTATACC TATAATTTTG
 22051 CTTGATTTTT TGTGTTCACT CATGTTTGGT ACCTAATCTC TAGGGGTTAT
 22101 TTCTATTCTT TATTTTCTTT GGGAGTAGGA GCTTTAGCGA CTTTAACTTT
 22151 AGCTACCCGA ATCGGGCGTT CTCCTATTTT ATATCCCTTT GCAAACCTCT
 22201 CTAAATCGT CCCCTCAGGA ACTTCAGAAG TCTCTTCTGT TTGCACCGCT
 22251 TCGTGTAGGA AGGGGTAAA CTTTTGGCCT ATTGAAGAAT ATTCAATAAT
 22301 ACCTTTTTCC TCGAAGATTT GTTTGAATTG GTTGAGAATC ATGTTGAATC
 22351 CGAGGGCCCA ATTTTTTACA TCGTCGGACA TTTGTGTAGC AAATCCGAGG
 22401 GCTTTCTCCA TGCTTCTAT GGGATTGAGA AAGTCTATTA AAGTATTTTC
 22451 TAAAGCATAC TGCATAAGTT CTTGGCGTTC TTTTGTAAAG CGTTTTCTAG
 22501 AATTCTCAGA TTCTGCTAGA GCCATGAGAT ACTTATCGTT TTTTCTTTT
 22551 AATTCGGTTT TTAGGGTGAC GATTTCTGT TGCAAATGTT CAACTTCATT
 22601 TTCGTTTTGA ACATTGCTTT CGTGTTGTTT CTCATTTTCA GGTGGGGTAT
 22651 CTGTCATAAC GTCTCCTTAG AGGGTAATAG TTTTATAGAA GAGTACTCCG
 22701 TTCTTAAAAT AGGCTCATTC GAAAGCTTAC AGTTAGAGGT GAGTGGTCTT
 22751 CTGAAGGATA GTTTAAATTT GTAGAACTT TGTGTCAGGG TTTCAATTAT
 22801 TTTATTCGCA AATAGTTTGA GCAAAGGAAG AGCTTCCTTA TAAGGAAGAT
 22851 TGATCGGGCC TAGGATACCT AAAGCTCCGA GTGGAGAGCG ATTCATATAA

22901 TAGGGAATAG TAATTACAGA ACATCCTGGA TTCGAGGTCC CTAAAATATC
 22951 AGAAAGCTCC TTCCCTATGA ACGCTGTAGC TCTTCCTTTA TGCATTCCTA
 23001 TATTTAGAAG CTCACACATT TGTCTGCGAT TTTCAAAAAG AGAGAGTCCT
 23051 AGAGCTAGAA CTTCAGGATC TTAAACGCT TCGTATTTC A GTAGTTTCGA
 23101 CATTCCTGTT TGATAGAGAT CTTCTTCACT AAAGTTGCAG TAGCGTGTTA
 23151 GATAGCGGAC AACCACCTCA TTATAGAGGG ACATGCTCAG GTGTTCTTCT
 23201 TTTTTCGAAA GTTCCTCATT TGTGGGGAGC TTTCCGATGT AGTTCTGCAG
 23251 GAATTTTTCT ATACGTTTGA TAGAAAGAGT ATCGCAAGCT TCAGGCAGCC
 23301 ATAGGGTGTC TGTGAAGATC TGACCAAAC CCGTAGAGAG GATGGTGACA
 23351 GCTCTTTGCT TATCGACCTG TGTAATTTGA ATATTGGTTA CGGAATCATT
 23401 TTCAAAGCGT GGGGAAGAAA AAAACGTAGG CAGGTCTAGG ATTTCTCCAA
 23451 GAAGTTCCGT AGCTTTTTGT AGATCCTTGA TAATATTGCG ACTTTCGCTA
 23501 GGAAGCTGAC TGATCTTATC AAAAATGGGG GCAGAAATCT CAGCTTCTGG
 23551 GCATTCTTCT TGGTGATCTA CATAGTGACG TAATGCTAGG TCTGTAGGGA
 23601 TTCTTCCTCC GGAAGTATGA TTTTTTTTTA AGAATCCTTC AGCTTCAAGT
 23651 TCTGCAAAGT AATTTCTTAT AGTTGCCGTA CTCAAATCAG AGCAAAAAC
 23701 TTCCTTTAAA GTTTTAGACC CTACAGGCTG CCCTGTTTTT AGGTACAAC
 23751 CTGTTGTAGC AAACAGGATA TCAAGGATTT TTGAATCTCG CTTTGAGACT
 23801 TTGGATCTAG CCATCTCGAG TCCTACTAGA ACAATCGTAA CTGAGAGCAT
 23851 TATAGGAAAA GAACCCGAGA AGGTCAAGAG AATTTAGCAC TCGAATTGAA
 23901 TGAATGCTAA CTTTTTTTACG AGGAGGGCAG CGATCAAAGA GCTAGGCTAA
 23951 GTGATTCTGA CACCAAGTAG GGAAGGCCTC CGGGGAGACT GTATACTTTT
 24001 CTCCAGATCG GGATTCAATT TCGAATATTC CCGAAGATTG GTAGGACTTT
 24051 CCTAAAATAA GCTTATAAGG AATGCCGATA AGGTCACTGT CTTTAAGTTT
 24101 AAATCCGAGT CTTTCATCTC GATCATCAAG AAGGGGCTCA TAGCCTTGAC
 24151 TTTGTAGCTC ATGATAAATA GTTCCGCAA GCTCTTGAGA TACAGTGTCT
 24201 CCTCCGTTAA AGGCGATAGT GATAGAGAAG GGAGCGAGTG CTTTGGCCA

24251 AACAAATACCA CGGTCGTCGG CAAGCTGTTC TACACAAGCG GCTAATGTTC
 24301 TTCCGACTCC AATGCCGTAG GTCCCCATCC AGCACTGCTG GGTTTGCCCCG
 24351 TGTTCATCTT GGAAGTTTAC CTCAAAACTA TCGGTATAGC GTGTCCCAG
 24401 ATTGAAAATA TGAGCAACTT CTATGCCTTG ATAAATGCGG TAAGGATGGC
 24451 CAGGATTTTC AGGACATGTG TCTCCCTCTT CAGCGAGTAG AAAGTCACCG
 24501 TATTGGGGGG GGAGGAGGTC GCGATCCCAG TTTACATTTA CGTAGTGCTT
 24551 ATCTTTAGCA TTGCCCCGAC AAACAAAGTT CGTCATTGGG GACGTTGTTT
 24601 CGTCTGCGAA AAAGTCTATG GGACAGTTTA GGGGACCGAT GAATCCTTTT
 24651 TCTGTGCCTA GAACGCGTTC GATTTCTTCA TCAGAAGCTA GAGCAATATC
 24701 ATCGGCATTC AGTTTGGAAG CGACCTTCAC TAGGTTGACT TGCCGATCTC
 24751 CTCTCATTC AATGGCAATG AATTTTCTT CATTTGAGTA GGAGAGTTTT
 24801 ACGACAAGGG TTTTAAAT TTTATGTAAG GGGATAGAGA AGAAGTTTGC
 24851 TAGAGCTTCT ATTGTTGTAA TCCCAGGGGT GGCCACTTCT TCGACGGGAA
 24901 GAAACTCGCG ATCGTAGGCA TGCTGTGGAG GAATGGAGAC AGCAGCCTCA
 24951 ATATTAGCTC CATAGGAACC GCTGACGCAG ATCGTGTCTT CGCCTAGAGA
 25001 GCAAAGGACC TGAAATTCCT CAGACTTTCC TTTGCCGATT TTCCCTCCAT
 25051 CAGCTGTAAC GATGACATAG GCAAGACCGA GACGATCAAA GATCTTACTA
 25101 TACGCAGAGC GGAGTTTTTC ATATTGCTCG TTCATTTGTT CGGGAGAGTC
 25151 TGAGAAGGTA TAGCTGTCTT CCATAAGGAG CTCTCGAGAG CGAATGAGAC
 25201 CGAATCGAGG GCGAATCTCG TCTCGGAATT TTGTAGCAAT TTGGTAAAGG
 25251 TGGAGAGGAA GTTGTCTTTT TGAGGAGAGC CATTGTGCAA CAAAAGAGCA
 25301 GATGACCTCT TCATGTGTAG GAGCTAGGCA ATGAGATTTT CCTTCGCGGT
 25351 CTTTGAGAGT GTAGAGCAGT CCTTCCGAAG TAAATGCCTC CCATCTCCCT
 25401 GTATGTTGCC AAAGTTCAGC ATTGTGGAGA AGTGGGAGTA GAAGTTCTTG
 25451 ACCTCCAATC GCATTAAAGT CCTCTCTAAT GATGTTTCATC ATCTTGAGAG
 25501 CCACGCGCCA TAACAGGGGT GTATAGGTAT AGACTCCTTT ACTTACTTTA
 25551 AATAGGTATC CTGCCTTTTC TAGGAGCTCG TTTGAGAGCA CAGCAGCGCT

25601 TTTATTTGCA TTTTTTGAAG TCTTATAAAA GAGTTGAGAC GTTTTCATAG
 25651 AGTGGGGCTG TTATCTTCAG TGATTGATCG TCAAGAGTCT ATAGAAAAAA
 25701 TTACAGGCTA TTGCGATTTT ATATCGAGAG TCTTTTAAGC TAGTAGATTT
 25751 AGTTTAGCTA TGAAGGAGGT ATCTTAACAC AGGGATTTTT CTAGATCAAT
 25801 GTTTTGCGTT GGGTTCTTAA CTTAACAATT GGGGTCAGGA TTATGTATTA
 25851 TTTTAATTAT ATTTTTGTTT TTATTTAGAA ATTTAATAAT CTCAACTTAT
 25901 AATTAAGTGT ATACTTATTA ATCTTTTATT TTTGTAATTG TAGTACTATG
 25951 TCGTCAGTAA ATCAAAGCTC TGGAACCCCG AATCCAGAAG AGGTAAC TTC
 26001 TCCTGAATCT ACGGAAGAAA AAAAAATGT TGTTCCTTCA GATGAGGCGC
 26051 AAGCCACGCA TGCTGTGGCT CTTCTTATAG TCACTCAACT TTCTCTTCCT
 26101 GAAGGTGTGG GGACCTCATC TGAAGAAACG GCGAGTAATC CGAGGGTAGA
 26151 CGAGATTGTA GCTGAAGTTT CTTGAGTCG GCGGTTGCT GATCAGATCT
 26201 CATCACTTGT AGAGCGTGTT GGAGAGCTTT TAGACGACCT TAAGGGTGCC
 26251 CAGTCCCTTT TCACTAGCTT TCAGTCAGAG TTGAAAACT GTCTCCGGC
 26301 ATGGAAATCT TCAACGAGAA GACTCGAAAC TCGAGGTGCT GGGGATAATG
 26351 CGGATATAGC GAGGCTGGAA TTATTTTCGTA GCGATTACGA GGCTGTCTTA
 26401 GGCCATGCGA ACCAGTTTCA TGGGAAGGCT CATCTCATTT TAAGTAAGTT
 26451 AACAGATGTA CATCACAAGC TACAGGGACT CAGTCGTGAA GATCTTTCCC
 26501 TGGCGTTTGA CAATAATGAT AGGGTTCTTG AGCATCTGGG TTCGTTAGGG
 26551 CTTGATGTAG ACGCTGAAGG TAATTGGTCT CTTTCTTG TG AGAGGGGGAT
 26601 TCCGCGACTG GTGCTTACTG CTGACAGTAT GCTTGTCCAG ATCAAGAAAAG
 26651 TGAATCTACC TACTGTAGAA GAATTGCGGA CTCTTCAGGG AACAAACGAA
 26701 TCTTCGTCTG ATCCTAGGGT TGAAGAAAGT TTGTCTTGCT GTGAAAGATT
 26751 GCTCAATGAA TTACGTCGTC TTTGGGCGAA TTTTGTAGGT TTTATTTTGA
 26801 GTTGCTATGA CAACATCGTG TTTGTTTTGA TGTGGATAGT GAGACGATT
 26851 AACCTTTTGC CTGGGCTGGG GTGTTTGCTT TTCCATAATC CCGATGCTTC
 26901 TCAAGAAGAC CAGAGGTCTT CTTCCGAGA GCGTTCTACA AGGAGAGAAC

26951 GCCTTTCTCG GCGATCTGAC TTATCTGAGG AAGAGATGAT TGTGAGAGCT
 27001 GAGGGAGAGT CTATACATCC TGAATCTCCC CATGGAGATG GCCGTAACCA
 27051 ACCTAGTCGA GGTGATAAGC AAGACTCTGA TAGTGAGGAA GAGACGGAGT
 27101 TATAATAAAG AATGATCTCT TCTGAGTAGA AACTATCGAA TCAGAATTGC
 27151 CAGCTTTTAC TCTCACCCAG TCCTATTTTT ATAGTTCAA TTGCGGTGGT
 27201 TTGTTTATTG TTTTTTTGAA TTTTGTAT ATGGTGGCGT CTAATATTGC
 27251 CGATTAAGGA GGGCGCCCTT TATGAATAGA AGAAAAGCAA GATGGGTAGT
 27301 GGCATTGTTC GCAATGACGG CGCTCATTTT TGTGGGTGT TGTCTTGGT
 27351 CACAAGCGAA ATCAAGATGT TCTATTGATA AGTATATTCC TGTAGTCAAT
 27401 CGTTTACTAG AAGTTTGTGG ACTTCCTGAA GCTGAGAATG TTGAGGATTT
 27451 AATCGAGTCC TCGTCTGCTT GGGTACTGAC TCCTGAAGAA CGTTTTTCTG
 27501 GAGAGTTAGT CTCTATCTGT CAGGTTAAAG ATGAGCATGC TTTCTATAAC
 27551 GATTTGTCTT TATTACATAT GACTCAGGCT GTGCCTTCGT ATTCTGCAAC
 27601 GTATGATTGT GCTGTAGTTT TTGGCGGGCC TTTGCCAGCG CTACGTCAGC
 27651 GCTTAGATTT TTTGGTGCGA GAGTGGCAGC GTGGCGTGCG CTTTAAGAAA
 27701 ATCGTTTTTC TATGTGGAGA GCGAGGGCGC TATCAGTCTA TTGAAGAACA
 27751 AGAGCATTTT TTTGATTCTC GGTACAATCC TTTCCCTACT GAAGAGAACT
 27801 GGAATCTGG TAACCGAGTT ACTCCCTCTT CTGAAGAAGA GATTGCCAAA
 27851 TTTGTTTGA TGCAAATGCT TTTACCTAGA GCATGGCGAG ATAGTACTTC
 27901 AGGAGTCAGA GTGACATTTT TTCTAGCAAA GCCAGAGGAA AATCGTGTGG
 27951 TTGCGAATCG TAAGGACACC TTACTTTTAT TCCGTTCTTA TCAAGAAGCG
 28001 TTTCCGGGAC GCGTGTTATT TGTAAGTAGT CAACCCTTTA TCGGTTTAGA
 28051 TGCTTGCAAG GTCGGGCAGT TTTCAAAGG GGAAAGCTAT GATCTTGCTG
 28101 GACCTGGATT TGCTCAAGGA GTCTTGAAGT ATCATTGGGC TCCAAGGATT
 28151 TGTCTACATA CTTTAGCGGA ATGGTTAAAG GAAACGAACG GCTGCTTAAA
 28201 TATTTACAGAG GGTTGTTTTG GATGATTCAT GGATCTTAGA GGTTAAAGTC
 28251 ACTCCAAAAG CCAAAGAGAA CAAAATTGTA GGCTTTGATG GACAAGCTTT

28301 GAAGGTCCGT GTTACCGAAC CCCAGAAAA GGGTAAGGCC AATGATGCTG
 28351 TAATTTCTTT ATTAGCAAAA GCTTTATCCT TACCGAAGCG TGATGTCACT
 28401 TTAATTGCAG GAGAACTTC TCGAAAGAAA AAGTTTCTTC TTCCTAACAG
 28451 AGTTCAAGAC ATTATTTTTT CTTTGCATAT AGACGTATAG CCTAACTCAC
 28501 AGCTACGTTT TTCCCTTTCT CAGATTTTTC CAATTTTTTG CTTTTAAAAG
 28551 ATAAGAACTG GTTGCGTTCT GTTTTATGAA GCTTGATTCC TAAGTACTTG
 28601 ATGATATCTT CATTAAAGGT GGTGTAGGT GACAGGCGTT GAGCAATGAT
 28651 TTTACGCAGG CTGTCCACAT CGTGATTGTT ATAGAGTAGG TGGTGCACAA
 28701 TTTTTCGAT TTGTTTCCT GATTTTTTGT AATCCACGCT ACAGGCAATG
 28751 CAGGCTCCTT CGGAAATTAA GGAGGTATCG TCGGTAATGA TAGGGATTTT
 28801 CTCTTTGAGG ATTCCTGAA GGAATGCGGT GCCTTCTTTA TGAGAAAGTG
 28851 GGGAGAGGGG AATGAAGATA GCTGAGGGGC GCTTGTCGAT AGCCTGGCGT
 28901 ATCCGGGTTT TGAATGTACT GCTGTAAATA GAGATCTCAA TGACCTCAAT
 28951 TCCTGAAGCA TGGAGTTTCT TAACAATTC TTTTGGAGA TCTGAGGGGA
 29001 AAGGTTCCGA GGGTTTTAAA TACACGATAG ATTGTGCATT GGTAGCTACG
 29051 GCTTGTATAG CAAAGCAGTA TTGATTGATG TCTAGAGTGT CATTCACTCC
 29101 GTAGATATTC ATTGTGTTTT TAGGAAGGGT TAGGCTTTCG CGATCAGGAA
 29151 CAGCGGCATA GATCACAGGT TTCTGTGTTT CAATGTGGCT CATGACCTTC
 29201 GTAGCAATAG TTCCTAAGGT GACAATCGCC ACGACATTTT TATCGGTATG
 29251 TAAGGAGCGA GCAATTTTCC TAGCCTTTAC GATACTGTCT TCAGCATTTA
 29301 GGACAACAAT TTCAGGAAGG TTCTCAAAT CTTTCAAGGT TTCTATACAG
 29351 CTTTTACTGC AATCTTCTAA TAGGGGATGG GGAAAGGATA AGAAAATTGC
 29401 GATTTTAGGA GAGGAGACGC TATCTGGTTG AGAACCACAA GTGGCTACAT
 29451 AGATGAAAGA GCAAAACAGA GAAAAGAAA ATAAGTACTG AGAGAGTTTA
 29501 CGTAGAATTG TCATGCAAGA ACCATCGTTT CTTTGTAGTA AGCGGTACAG
 29551 AGGCGGTCAC AGGCTAAAGC GATATTTTGT GGTGTGTCA GAGCGGAGAA
 29601 ACGAACAAT CCTTGTCCAC AGGAACCAA ACCGTGGCCG GGAGTCACTG

29651 CAATATGATA CTGATGTAAG AAGAAATCAA ACGCTTCTTC ATCAGAGATT
 29701 CCTTCAGGGA GTTCTACCCA AAGGTAAGGG GCATGATCGC CACCATGAAC
 29751 TGAGAATCCT GCAGTTTCTA AGCTTTTTTT AAGTTTCTGA GCATTGGTTA
 29801 GATATAAAGA GATGGCGGGA GGTGTCGGAA ATAAATCTAG GCCGTAATAC
 29851 CCTGCTTCTT GCATGAGGAG AGATGCTCCG TTAAATGTAG TCGCAAAGAG
 29901 CCGTTTCCAA TCGTTGATCA TAGGTTTCGT ATTGTCATAG GTGAGTTCTT
 29951 TAGGGATCAC GTTCCAGGCA AGGCGCATGC CAGTAAAGCC TAATGATTTA
 30001 GAGAAAGAGT TGATTTCTAT AGCACAATAT TTTGCTTCAG GGATTTTCGAA
 30051 GATGCTTTTA GGTAGGCTAG GATCTGAGAC AAAGGCGCTA TAGGCCGCAT
 30101 CAAAAATAAG AACGGTTCGG TGCTGATTCT CGTAGTTCAC AAGTGCTTGG
 30151 AGTTGTTGAA AGGTTAGAAC TGTTCCTGTG GGGTTGTTAG GATAGCATAG
 30201 ACAAAGAATG TCTAGGGATT GTTGGTTCGG AAGTTCTGGA ATAAACCCAG
 30251 TTTCTTTTCT GCATGCTAGG GGGATAATGT CGCGGATTCC TGTAATGTGG
 30301 GCAATGTCTC TATAAGCTGG ATAGACAGGA TCCTGTAGAC CTAGAGTCTT
 30351 TTCTGAGCCA AAAAAAGAAA AGAGACGGAA GATATCAGGT TTGGCACCAT
 30401 CCGAAATAAA AATCTCTTCA GGGGAGATTC TATTTTCATA GACTTCAGAG
 30451 GCAATTTTTG TGCCTAATTT TTCTAATCCG GTTTCTGGGC CGTACCCACG
 30501 ATAGGTCTCT TGTTCCTCTT GAGAAACGCA GAACTCTTTG ATTGCCTGAG
 30551 TAATAGAGCG GCAGAGAGGT TGTGTCGTAT CTCCGATAGA AAGATCTATG
 30601 ACAGAGATTT CTGGATTCTC CTTGCGAAAC TGAGCAAGCT TTTTACTAAT
 30651 TTCAGAAAAT AGATACTGAG GCTTGAGAAG AGAAAAGTGG GGATTTCTAC
 30701 GCATAGCGTG CTCCAGGTAA GGGCGTGATT CACTTGGAAG ACATCTTAAG
 30751 AAAGCCCCAG CTTTTTGGAT AGCCATTTTC TGATTTTCTT TTAACCGCCT
 30801 CTATAAAGGC AAGGAGTGCT CTGTAGTTTC GGTTAATTAT TAAGATATAT
 30851 TTATAATTCT GTTTATTAAA AGTTTTTTAA ATCTTTTCTA ATTTGCTCAC
 30901 TATAATTAAA GGATAAGATT TGAAAAAATT TTTTAGGTAA TTATGATAAG
 30951 GGTCAATCCT TATGGAAGTT ATAGGGGTAG GAATCCTTCT CCAGAAGATG

31001 GGAAAAAGGA TGTACCCCTT TCAGGGAAC CTGCTTGCA TCGTCGTGGT
 31051 GGGATTCGTA GAAAGCATAA GAGTGCTTCA GTTGGGGTGA CCTCGGGTTC
 31101 TAAGACGGGG AAAGCTTCTT TAGAGAAGAA GGTCAAAGGC ATTTCAGAAG
 31151 CCCATTTCAA ATAATCCAAG ACAGAAGGTT CTCATTCGAA GACAAGCAAA
 31201 GGATTCGTTG GCAGATTTGT TCAATGGATT AGAACGTTTA CAGGACGTGG
 31251 AAGCAAGAAG CGTTCTCCCT CAAGTTTTTC TCCAACGCAC CCTTACATAC
 31301 GTTTGCGAAC TTACACACGC AGTCCAAAAC AGAGTGGTGT AGAGAGAAAA
 31351 CAAGAAGATG CTGAGACCTC ATTTATAGAG ACACCCAAAG GGATCTTGAA
 31401 AAAGCCTGGA AACAAAGACC CCAAAGGCAA GCACGTCCAT TGGAAAGACA
 31451 GCTAATCCGG ATCAGAGCTG AGATCCCTCC TTCTCTTCAT TAAGAGGGTT
 31501 CAGAGTTTTA GGTCTCTGCG GTAGAAGGA GGAATCCAT AAGGAGAAGA
 31551 CCAGTGGATT TTCAACGGAG ATTCTAGAAG GAAACAAGCT GATTCTACAT
 31601 GGAGATGTTT GTAGGTCTGT TGTGAATTCC TCTCTTACA GATCGGAAAG
 31651 ATCAATGAAA AGAGTTAGGG TCTCTTAGTT TTAAGAATTG CTATACAATA
 31701 TTGCGAAATA AATACTTTCT TCTCTATACA CTGTCTTTAC GATGAAGACA
 31751 GCTTTTCACT CTGCTATTC TTGGTTTTGT TGGCTCTTTA GCTTCTTGGT
 31801 ACTCTTTGTG GGTGGCATCG CTGGGGGAGA GCCTTTGTGC CCCGATTGCA
 31851 AATACGAAAC TAAGTCTGTT TTACGTTCGG ATCAGCTGCC GGATCATCTC
 31901 TGGAACATG AAAACGACTG TTATCTTACA GGTTATGTGC AGTCTCTTTT
 31951 GGACATGCAT TTTTLAGATA GCCGTACGCA AGTTGTTATT GAGAAGAATA
 32001 GAGCGTATCT TTTCTCTTG CCTGTAGATT CGAGTTTATC AGAAGCCATT
 32051 ACCAACTTTG TTAGGGATCT TCCCTTCATA TGTGCTGTGG AGATTTGCGA
 32101 GCGTCCTTAT GGTGAATGCA TAACGAGATC TTCTGCGGAG CGTCCCTTAC
 32151 TCCCTAAAGA GAAACTTTA GGAATGCCAA TTTTCTGCGG CAAAGAAGGG
 32201 GTATGGTTAC CTCAAATAC CATTTTGTTT TCTCCTTTGA TTGCAGATCC
 32251 TCGTCAGGTT ACCAACAGTG CTGGCATTCG TTTTAATGAG AAGGTCGTGG
 32301 GGAATCGGGT AGGTGCTACC ATCTTTGGGG GAGATTTTAT TCTCCTGCGT

32351 CTTTTTGATG TTTCTCGATT CCATGTAGAT TGTGATTTCG GAATTCAAGG
 32401 AGGAGTCTTC TCAGTTTTTG ATTTAGATCA TCCTGAATCG TGCATGGTAA
 32451 ATTCAGATTT CTTTGTTGCC GGA CTCTGGT CAGGGGCTAT AGATAAATGG
 32501 AGTTTTAGGT TTCGATTGTG GCACCTCTCG TCCCATTAG GAGATGAGTT
 32551 TATTCTTACG CATCCAAATT TCCCAAGATT TAATTTGAGT GATGAGGGCG
 32601 TCGATCTCTT CATTTCTGTT CGTTACACAC CACAGATCCG CTTGTATGGC
 32651 GGCTGCGGTT ATATTGTAAG TAGGGATCTT ACTTTTCCTG AGCGGCCGTT
 32701 TTACTGTGAA TGGGGTGCGG AACTCAGACC TTTTGGTCTG AGAGAAGGAA
 32751 ATCTCCACGC ACAACCGATT TTCGCGATGC ATTTCCGTTG TTGGGAAGAA
 32801 CAGAAATTTG GCTTGATCA AAGCTATATT TTAGGCATGG AGTGGGCCAA
 32851 ATTTCAAGAA ATCGGAAGGA AAATCCGTGC TGTTTTAGAA TATCATCAGG
 32901 GATTTTCTAA AGAAGGCCAA TTCATTCTG AACCGTGTA TTA CTACGGT
 32951 TTCCGTCTTA CCTATGGATT CTAAACGATA ACAAACCAT CTT CAGGGAC
 33001 AGGGTGTTCA GGTCTATGG GCAGCGTATG ATTGAGATAT CTGCGGAAGA
 33051 TTGCAATGCC TGCCTTGCT GAGGATAGGC AGAATAGGCA GTTGTGTACC
 33101 CATTCAGAAC CTCGAATGGT TCCTACAGCA TGATTGGAAT TATACAAAGC
 33151 TGTGATCTTG GATTTCCAAT AGTCTACAGG TCCGATTAGG AAGACGGGAA
 33201 CAAGAGCTTT TTTCCCTGTT TTGAGACTAA TAAGCTCCAG AAGGAGTTCG
 33251 AAATCGGTTT CCATGCCTCC GATAACAAAT ACAGCAAGGT CGACATGGAA
 33301 GTCGGCCTGA CGTTCTAAAA GATCAGGAAT AGCATAGCTC ATTTTAGCTT
 33351 CTACATAGGC ATTCGTGGTA TCCAAGCTAA TTAGATTCCC ACAAGAGAGT
 33401 ATGGAGAGTT CTGTAGCTAC ACGATTCTCG AGTTCCATAG CTCCAGAACC
 33451 CCCTCCTGTA AGGATTGCTA ACGGTGTCTG TGGTGGAAT TCTGGGATCG
 33501 TGAATTGCTG AGAAAGAGTA TGCATTCCTG TCAGGAGCTC ACGGAGAAAC
 33551 TCATCATAAT CCCAGCGAT TAGGCAGGAA CCATGAATTC CTATAAAGTA
 33601 GGATTGAGCA AACTGTTT CTTGATGTTT AGGGACAAAC ATGCCACAT
 33651 CTTTATTTCT GCGTTTGATG TATTGTAAGA GTCGTTTCGA TTCTAAGTCT

33701 GCCCAAAATA CAGAAATTCC TGCAAAATAT AGATCGAGAA GGAAAGAGCG
 33751 ATCTCGATTC GAGAAAAACT CTCCAGAAGT GGGAGAGGGA ATCTGAAAAT
 33801 AGATATGTTG CAGGTAATAG CGAGAGTAGT TAGAGAGGAA CATGCCCTTC
 33851 AGCGATGCTG AAGGGAAGTA GCGGGAAAAT AAAACTCCTT GGCTTGTGAT
 33901 ATGATCTGTT TCCATGGCTT TTAAAAAAGG GAAACAAGGT TGGTCTTCAA
 33951 TGTGCTTTTG AATTTCCCTA GCATGTCTTT CATCTGATGG GGAGATTCGA
 34001 GGT TTGATGA TCCAAGAGTC TTGGGAGAGC TCAAGCAGCT CACTACCTTT
 34051 GGAGATAAAC ATCGCAGCTT GATCTTCGCC TTCCGGTATG GATTCAAAAA
 34101 CACGAAATAC CTCTTGAGGA GATTCTAAGG TTTCTGGAG CATATCTCTA
 34151 TAGAAGAAAA ACGAATGCTC TTTGTAAGGC TCAAGAGTAA AAAATTCTAA
 34201 AGGTATTCTC TCAATAGGTT CTGAAGTGCT GCCGTAGAAT TCATAAATAT
 34251 CTCCAGATTC TTGTGTGGTA GGTTCGAGAA TATCCGCTGC GGTGTGACGA
 34301 AGCCCTTGGG GGAGTAAGTC CTGAACGACT CTTGCAAATA CGGTTCGGAT
 34351 GTGCAGAGGC TCTGTCTTTA TGAGAAGAAT TTTATGATCT TCGGGAACGG
 34401 GAGGACGATC TGTTACCATT TGATACAAAG GAAGAACTT ACGTATTTTT
 34451 AAATGGGGAC GCGTGAGTGA TTTGCTCATT AAGGGAAGGA ACCCATAAAT
 34501 TGTCTCTTCG TAACAGATTG TTCCTGGAAG GATCGGAAGG AAGACAACAA
 34551 GCCGATCATT AATGATCTCT AGAGTGATGA AGTGCTCAAG TTTTTTCCCA
 34601 AAGCGTAGGA GCGGAGATCC TGTACGGTCT GTGTGCGTAA ACATCCTGTT
 34651 GAGATAACAA GGCGAACGTA CGAGTCGGCG ATCATCAGCA GCAAAGAGCT
 34701 TGCAGACAAA ACTTCCAGGC TCTAGGAGCT CCAACATAGC AGTGGCTATA
 34751 GGATCTTGGC TCATGAAGAG AACGTGTAGA CGAGCTTCTT TTCGGGCTTT
 34801 ATTTAGCTCC AAGTGGTTTA AAACGGCTTC GACACCTAGT TGGGCTAAGG
 34851 AACTTTTTTA ATTTACTTGT ATACACTGTT GAGGCAGATG AAATCCAAGA
 34901 AAGTACGCAG GAATATTCTC AATGAGGACC TCTCCTTCGT AAATGTGGGG
 34951 CGAGAGTTTT TTCAAATGGG AAACGAGTCG TCCGTCTGGG GAGGCTGCAT
 35001 CATGATGCGC GTGGAGTAGG TTATACATAA TTCATCGTAT TTAAAGTTA

35051 AAAGCAATAG GTAAGGTCCC TCTTTGCGCT TCTAAGACTA GCGTTTCCAG
 35101 AAGAATTTCC TACAAGGACT TCATTCTGCT TGTGTGTATA ATTTTGTGCA
 35151 ATTTGTCAAG GAGGGGGTCT TGGAACGAGA AAATTGTGTT GTGGAAGGGA
 35201 ATTTTAGGAA AAAAGTTTCT TAGAAAATAG GGCAGAAACT TTCTCTTGGA
 35251 GAGATTGAGT AAAATGTAAA GAATAGTCTT TGCAATTGAG AATTATTTCT
 35301 CTCTGGTTAC AATGGAGGAT TGGCTAAGGA GGATAGTAGG TATGCAGATT
 35351 CCAAGAAGCA TTGGTACTCA CGATGGTTCT TTCCATGCGG ATGAGGTCAC
 35401 AGCGTGTGCT CTCCTTATTA TTTTCGATCT TGTGGATGAA AATAAAATTA
 35451 TACGCTCTCG AGATCCTGTC GTATTATCGA AATGTGAATA TGTTTGTGAT
 35501 GTCGGTGGTG TTTATTCTAT AGAAAACAAG CGTTTTGATC ATCATCAAGT
 35551 CTCTTATGAT GGATCTTGGA GTAGTGCAGG TATGATTCTG CATTATCTTA
 35601 AAGAGTTTGG TTATATGGAT TGTGAAGAAT ATCATTTCTT TAACAACACT
 35651 TTGGTACATG GTGTGGATGA ACAAGATAAT GGCAGATTCT TCTCTAAGGA
 35701 GGGATTTTGT TCGTTTTCTG ATATTATTAA AATTTATAAT CCTCGCGAGG
 35751 AAGAAGAAAC TAATTCGGAT GCGGATTTTT CTTGTGCTTT GCATTTTACC
 35801 ATCGACTTTT TGTGTCGGCT AAGGAAGAAG TTTCAGTATG ATCGAGTTTG
 35851 TAGGGGGATT GTCAGAGAAG CCATGGAAAC CGAGGATATG TGTTTATATT
 35901 TTGATCGTCC TTTAGCATGG CAAGAAAATT TCTTTTTTTT AGGGGGAGAG
 35951 AAGCACCCTG CAGCTTTTGT TTGTTTTCCT TCCTGCGATC AATGGATTTT
 36001 ACGAGGGATT CCTCCGAATT TAGATCGCCG TATGGACGTT CGTGTTCCTT
 36051 TCCCTGAGAA TTGGGCAGGT TTGTTAGGTA AAGAGTTGTC CAAAGTATCA
 36101 GGGATTCTTG GGGCTGTGTT CTGCCATAAA GGTCTTTTCC TTTCTGTATG
 36151 GACAAATAGA GAAAGTTGCC AACGTGCTTT GCGGTTAACG TTACAAGATC
 36201 GAGGGATCAT ATGACAGTAT TCAAACAAAT TATCGATGGA TTGATAGATT
 36251 GTGAAAAGGT ATTTGAAAAC GAAAATTTCA TAGCTATAAA AGATCGTTTT
 36301 CCTCAAGCTC CTGTTTCATCT TCTTATCATT CCTAAAAAAC CTATACCACG
 36351 ATTCAGGAT ATCCAGGGG ATGAGATGAT TTTAATGGCA GAGGCTGGAA

36401 AGATCGTGCA AGAGCTTGCT GCAGAATTTG GAATTGCCGA TGGGTATCGT
 36451 GTGGTTATCA ACAACGGTGC TGAAGGAGGA CAGGCGGTAT TTCACTTACA
 36501 TATTCATCTT TTAGGTGGGC GTCCTTTAGG TGCTATAGCC TAATTTTCTT
 36551 TTGTTTCTGT GGATCCTTGT TCGGCTCGGA GTCCCTCCGT TATCAATTGT
 36601 TGATCCAAGA TTTTGCAAAA GTTTCAGAAG AGGGCATAGG CCTTTTGGAG
 36651 TCTAAAGAGT ATTCTTTACT TCAGGCTAAG CTAGTTTAA GGGCTCTGGC
 36701 TCAAAATTCT TCTTTTGATG ATTGGTTTAG AAGTTTAAAG AAGTGTGAGA
 36751 TTTCTATCC AGAGTTAGCT CATGATCGCG ATGTCTTAGA AGAATTGGG
 36801 ATTCAAGTTC TCGGTGAGGG AATCGAAAAT CCTTCCGTGA CCGTTCGTGC
 36851 TGTGAGTGC CTTGCTATTG GGCTTGCTAG AGATTTTCGC TTGGTCCCTC
 36901 TCCTGCTCCA AAGTTGTAAT GATGACAGTG CTATTGTTG ATCTTTGGCT
 36951 CTTCAGGTTG CTGTGAACTA TGGCTCTGAA AGTTTAAAA AGGCCATTGT
 37001 AGAGCTTGCC CGTAATGATG ATTCTATTCA TGTTCCGATT ACAGCATATC
 37051 AGGTGGTCGC TCTTTTACAG ATAGAGGAGC TATTGCCATT TTTAAGAGAG
 37101 CGTGCTGAGA ACAAACCTGT AGATAGTGTA GAACGTCGAG AGGCGTGGA
 37151 GGCTTGCTTG GAACTCTCTT CTCAATTTCT AGAGACGGGT GTAGCTAAGG
 37201 ACGATATTGA TCAAGCGTTG TTTACTTGTTG AAGTGTGCG TAACGGTATG
 37251 TTGCCAGAGA CTAAGGAGAT TTTTACAGAA CTCTTATCTG TAGAGCATCC
 37301 TGAAGTGCAG GAGTCTCTCT TACTTTCTGC TTTAGCTTGG AGTCATCAGC
 37351 TACAGAATCA CAAAGAGTTT CTTAGTAAAG TCGGCCATGT GATGTGCACT
 37401 TCTCCATTTG CAAAAGTACG TTTTCAAGCT GCTGCACTTC TCCATCTGCA
 37451 TGGAGACCCT TTGGGCAGAG ACTCTCTGGT TGAGGGCTTG CGCTCTCCTC
 37501 AACCTCTTGT GTGTGAGGCA GCTTCGGCGG CTCTCTGCTC TTTAGGAATC
 37551 CATGGAGTCC CTTTGGCAAA GGAGCATTTG GAGAGCCTTT CTTCTCGAAA
 37601 GGCTGCTGCG AACCTCTCCA TTTTGCTTCT TGTGAGCCGT GAAGATATTG
 37651 AAAGAGCTGG AGATGTGATT GCTCGCTACC TCTCCAATCC TGAAATGTGC
 37701 TGGGCTATAG AGTATTTCTT ATGGGATGCA CAATGGAATT TACGTGGTGA

37751 TACCTTCCCT CTATATTCGG ATATGATTAA ACGTGAGATT GGTAGGAAGC
 37801 TCATTGCGCT TTTGGCAGTA GCTCGCTATA GCCAAGCCAA GGCTGTAACA
 37851 GCAACGTTCC TTTCAGGACA GCAAGCTCAG GGATGGAGCT TTTTTTCTGG
 37901 AATGTTCTGG GAAGAGGGAG ATGTGAAAAC TTCTGAGGAT TTGGTTACAG
 37951 ATGCTTGCTT TGCAGCAAAG TTGGAAGGAG CGTTAGCCTC GCTATGTCAG
 38001 AAAAAAGATC AAGCTTCCCT ACAGAGGGTC TCTCAACTTT ATAATGACAG
 38051 CCGTTGGCAA GATAAATTAG CAATCTTAGA GAGCGTTGCT TTTTCTGAGA
 38101 ATCTTGATGC TGTGCCTTTT CTTCTAGACT GCTGCCATCA CGAAGCTCCT
 38151 TCGCTGCGAA GTGCAGCAGC GGGTGCTCTT TTCTCTATTT TCAAATAAAT
 38201 ATTAATAAAA TTATTCAAGA TATAGAAGAA AAACCACGCA GTTGTAATTT
 38251 CTATTTCTTA AAAATAATTT TTCAGACTGA CTTTATTCTT TCATTTTTAA
 38301 GTCTTTGGAG ATAGAAAACT TTGTTATAGA TTTTATCTG GTAGCTTTTA
 38351 TAATTTATGA AGAGCGTAAG CTCAGAGCCT GTATGTCATG CACAACCTGT
 38401 ATATAAAATC AGATTTGTTT TTTGAATCTC TATTCTCGTT AAGATTTTCGT
 38451 TATTCTGGAC GTTATCTCCA TCACCACTCC TAATTTTCCT AGCATTTCTA
 38501 TCTTTAAGCT CAGCACCGTA GCTTGCTTAA AGGAAATATT TTTCATTTAG
 38551 GTTGTGGAGT TCTTTATTTT ATGAATTTT CATTATTTT ATTTTTCCTG
 38601 ATAGCTATTC AGGGAATCTG CTTGTACGTG GGACGTCGTG GTAGCAAAAA
 38651 GGTAAGAAGT CGCGAGAGCT ATTTTCTTGC AGGAAGGAGT TTAAAAATCT
 38701 TTCCTTTGAT GATGACATTC ATTGCCACCC AAATCGGTGG CGGTGTACTT
 38751 CTTGGGGCTG CTGAAGAGGC CTTCTGTTAT GGTATATGGG GGATTCTTTA
 38801 TCCTTTAGGA GTCGCTTTAG GGTGATTTT CTTAGGAATG GGGCCCGGA
 38851 AGCGGTGGC AGAGGGATCG TTAACGACCG TAGTCTCTAT CTTTGAAGTG
 38901 TTTTATGGTT CTAAAAAGCT CCGTAAGATC GCATTTTAT TATCCGCAGG
 38951 TTCCTTATTT TTCATCCTGG TCGCTCAGGT GATTGCTTTA GATCGGTTGT
 39001 TTAGCAGCTT CCCTTTTGGC AAGTACGTAA CCGTAGCATT TTGGATTGTC
 39051 TTAGCATCCT ATACCTCAAC AGGAGGGTTT CGCGGGGTCG TACGTACTGA

39101 TGTGATCCAA GCAGGATTTT TTCTTATTGC GGTGCTCGTC TGTGGTGTTC
 39151 CTGTATGGCT CTCTGTCCCT AAATCCTTGT CTGTGTTGGA TCCTTTCCAA
 39201 TCACTTCCTT GTGCGAAGCT TTCCAATTGG ATATTCATGC CTATGCTCTT
 39251 TATGCTTGTT GAGCAGGATA TGGTGCAAAG GTGTGTGGCT GCCTCCTCTC
 39301 CAAAACGCTT GCAATGGGCG GCTGTAGGCG CAGGCCTTGT TCTTCTTCTT
 39351 TTAACTTTA TCCCTTTATT TTTAGGTTCT TTAGGAGCTA AAGCAGGCCT
 39401 TAAAGCAGGA TGCCCTCTGA TTGATACCAT TGCATATTTT TGCAATCCCT
 39451 CACTAGCAGC TGTGATGGCT GCTGCCATCG GCGTTGCGAT TCTCTCTACC
 39501 GCGGACTCTC TTATGAATGC TGTAAGCCAG CTAATCGCTG AAGAATACCC
 39551 TACGTTGAAA GCCCCTTATT ATCGTTATTT AGTATTGGGT TTGGCGGTTG
 39601 CAGCTCCTCT TGTGCTATT GGTTTTACAA ACATCGTAGA TGTCTTGATT
 39651 TTAAGCTATA GCCTGTCAGT GTGTTGTCTT TCAGTCCCTG TGGGTTTCTA
 39701 TCTTCTAGCT CCTAAAGGTC GCCGTGTGAG CGGAGCTGCT GCTTGGGCAG
 39751 GAGTGCTCGT TGGTGCTCTG GGCTATGGAT GGGTTCAGAT AGTCTCTTTG
 39801 GGGATGTTTG GGGAGCTATT GGCTTGGGTA GGTTCCTCTAG TCGCCTTTTC
 39851 CTTTGTAGGA TTTATTGAGA TCACTTGGA AAACAAAGTC AAAACGCAAA
 39901 CTTAGATAAC CACTGCATGA GAAGATATAA CTAAAATAGA TCCTGAGTTG
 39951 TTTAGGTTTC TCTTAGATCT GATAGGTTGC GCTTAGTAAG AGATCGTCAG
 40001 TTTTTTAAGT TGTGTTTAGA ATCTGATACC TCTCCTTCTT TTCCAAGAAG
 40051 AAGAGGGGTT CGTTTTATTT TTTATTATTC ATTCGTAGGG GCGGGAAGCT
 40101 GTTTTAACTG ATAGAGCAGG TCGATAAGAG AGGAGAGCAG GGCTTGATAG
 40151 AGCGTCTCTT CAGAATGGTG AAGAGTGCCC TTAAGAAGCT TTCTGCAATA
 40201 CTTTAATAGC CAATACAGAG CGCAAAGTAA AACTGTAAA AGGTAAAATT
 40251 TACACGCTTG TTTTATCATA AACCTCCAAA GAAAGACGCG TGTTGGCGTT
 40301 CTTCCAATTG GCTGTTAGAT AGTGAAAATA GTATTTACTA TTCAATAAAA
 40351 ATATTTATAT TAAATATAAT AAAACAAAT TTCTAATAAA CTTTTTAAAA
 40401 GTATTTGGCT AAGTTTAATT TAAGAGTTCT CAAATAAAAG ATTTTTTAGT

40451 CTCTTTTATT TGAGAATCAT CAACCGATTT CAAGAAATCG CATGAGCCCC
 40501 TGATTCTGAT GGGCAGAAGT GCTTTTtagc AAATAGGTGA AGCTCCTCGG
 40551 GCGTGATCTG TTGCTTCGCG TCACAAAGAG CCTTTTCAGG GTGTGCATGC
 40601 ACTTCGATCA TCAGACCGTC GGCACCTACC GAGAGACCAG CAGAGGCGAG
 40651 AGGAAGAACT AGAGAACGCT TCCCCGCTGC GTGGGAAGGA TCTACAATTA
 40701 CAGGGAGAGA AGAGATCTCT TTAAGGAGAG CCACGGTATT GAGATCTAGC
 40751 GTGTAGCGCG TAGAGTGCTC AAAGGTACGA ATTCCTCGTT CACAAAGGAT
 40801 TACCCCAGGA CAGGAGGGAG AAGAAGCAAG GATGTACTCC GCTGCGCATA
 40851 GCCACTCTTC AAGAGTAGCT GCTGGACTGC GTTTTAGGAT AATCGGACGA
 40901 TGTGATTTGC TGACCTCTTG TAAAAGAGGG GTGTTATGCA TGTTTTTGGC
 40951 TCCGATACGG AGGATATCCA CATGTTCCGC AGTAATTTCA ACATCTCGGA
 41001 CATCTAAAAC TTCGGTTTCT GTAGGGAGAC CATGGATGCT CTGTGCTTCC
 41051 TTATGCCAAA GCACACACTC TTTCTCCCAT CCTTGAAACG AAAATGGGCT
 41101 TGTCCGTGGT TTTCTGATTG ATCCTCGGAA TACCTGAGCT CCTGCTTCTT
 41151 TAACTGTAAG AGCTGAAGAG ACTGTATGCT CGTAACTTTC TAAGGTGCAG
 41201 GGGCCTGCGA TCAGTATTGG CGATCCTTCT CCAAACGATA GATTGGAGA
 41251 AATAGGAACG GTATGGACCT CGTCAGGATG CTGTTTGAGG GTGCGCGGTA
 41301 GGGGATAGGT AAACGTAAGA ATAAGTACCT CATGCAAAAC TAGGGATTG
 41351 AGCTGTCCGG TTTTGAGTTC TGATCTTCAG TTCTAGGGTT TCTAGGCAAG
 41401 AGACAACCGT AGTGGTTGGG GAAGCGGATT AAGAAGATAC GAGCGTCTCT
 41451 TTCCGGAGAA ATTGTATTTA AAAATCCATG CATCGCAACA AAGTTTTCAA
 41501 TATTTTCTTC TGAGACAGTA TCTCGATCAG ACATGATCAG ACAAACAGCA
 41551 AACGGCTGCA AAGGTACAGT AGTCATAGCA GCGACGTCGT TAGCCTCAGC
 41601 TTGCTCTTTG ACGTCATCCA AGATATCACC GAGTTCATTA GAACTTAGGG
 41651 AGTGGAGGTA GCCTATCAGT GATTGTCTAA AACCAGAGCG ACAGAAACCA
 41701 TCAGATGCTG TGACCGTGTT TTTTAAGAGA TTACAAATCA ACTGATCTCG
 41751 AGCTCTAGAG TATTGATTTA TAGTCGCTGT AGCTAGAGGG AAACCGTATT

41801 TTAACGAAAG TAAAAGCTTA TCTTTTACTT GTGTAAATCT ATGAGAAGCC
 41851 TCTCTTTGGA AAGCGTCTAA TCGAGCAGCT AGCGGAGTCT TTTTTTGTAG
 41901 AAGTTTAGGG TAGTTGTAA GGAAACAAAA GATTAAAAAA GATTTGAAA
 41951 ACTGATCACT TAAAGCTGTG GGAGAAGAAG GCGAACCCAG TTTTGATCCG
 42001 AAGAGATTTA ACCGTGTAA AAAAACGTC CACCCATTCC TAAAGGTACT
 42051 TTCGATTTCT TTTTCTAAGA GTATGGCCTC GGTTTCCAGT TGATAATTTG
 42101 GGAAAGTAGC TCGTAGAAAC TGCATACAGT TCTTTTCAAG TATGGGATCT
 42151 AAAACCATCA CTTCAGGATG TCGTGACCAG AACTTGCAGA ACGTCTCTTT
 42201 ATCTTTATTA GAAGGGGTTG AAGGATAGCG TGCTGGAGGC TTGAGTACTA
 42251 GAGACCCTAG AATGGAACGT AGAAGTTTTT GGAGTTTCTT TTCGTTTTCT
 42301 ACTTTTTTAA AGGCATATTG TAGAAAAGGA AGTAGAAGCA TTTTAACTC
 42351 TTTAACGGCA TCTTTAGTTT CTGGACGATT ATTCACTTGA GGATGCTGTA
 42401 AAATAAGAA TAAGAGCTGT ACTGCCTGTG TATAGAGGAG GTCCTGCTCT
 42451 TGTGAGAGAC CACTGCCTTC ATCACGCGTT AAACGTAAAC TTCTTGCCGA
 42501 GCAGACAGTA TTGAAGTGCT CTTTAAATTC AAGCTTACTA TGAATTGCGA
 42551 CTAATAGCTT GTTTACAAGA TTTGAAAGTA AATCCGGCTG ATTGCTTAAA
 42601 GGAGGGGTTT CAAAGGATGT GGATATTACA TTCTTTAATA GGTTTGCTAT
 42651 GCGTTTATGT AGTGTTTCGA GTTGAGGCAT AGAGCTCTCT AACGTCAAAC
 42701 TCTCATGAGT CGTAAGGATA GTCATGATGT TAGAAGAAAT CACCTCTCGT
 42751 TGCCAAGAAG AATCTAAATC TTCATTAAGG TAGCTGAGGG AGGCGTTACT
 42801 TATTTGTAAG ATCTCTTGTA TAGAGCGTTG GGGGGGCGCG GACGCTATAA
 42851 TGATATCCTC TAATGATTTG TAGAAAGTAG CGCAGAGTTT GTGAAATCTT
 42901 ACAGAGGGGC AGGTACATAG AAAAGCAAAG AAAGAAAGAG CTGGAGACCA
 42951 AGGGAGATCC CCAGATCGTC TGTTTTCTAT AGTTGCAAGA AAGCGAGAGT
 43001 ATTGACATCG GATTGCTTCA GGGAGGTGAT CTTTGA CTAT AGCATTGAAT
 43051 CTCTCAGTAT CCCACCCTGA ATCAGCAATG CGTCGACTCA GCTTAGCAAA
 43101 AGCCTCTCGT ATTTCTCTTT CATATTGCTT TCGAAGTTGT TGATCTTCTG

43151 GAGACATCTT TTGAAGCTCT GCATCAGTAA GAAAGAGAGC GGGCAGATGT
43201 TCCAAGAGAA GAAAGAGTTT GTCTCTAGAT TGCATGTAAG GAAGTGCGGA
43251 ATGATCAATG AGATTTTGGG ACAATTCCGA ATAGGGACGA TTTGCAGCAA
43301 GAAATTCCAG AACACTGGGG AGGAGATCGT CCTGCATATC AGAAATAAAT
43351 AAGGCTCTCG CTGTTTCTTC ATTACTTGAA GCTGCGATTT GTTGGCGAAT
43401 CGCATATGCA GAGAGTTTTC TTAGAAAGGC TATCAGAGTT GCAGTGTGTT
43451 TCTGAGAAAG AATCACCCCG TCATAGAGGT CTATGAAGGA GCAATAAGTA
43501 CTGCATAACT GAAGGAGTTC AGCCATTTCT GCACAAAGAT TCGCATTTGC
43551 CGGAGAAGAG GAGGCGAAAG GCAGGTCAAG GAGACGTGTG GCTTCCTGCT
43601 CAAAGACAAT ATCATTTCTG CTGCTCTCTT CGTAGAGAGC AGATAGCCAT
43651 CCTACAGCAT AGGCACGATA AAAGCAGTTC CCATCTCCCG GTACATTCAC
43701 AAGGTAGTAA TTGTCATTTA GATAGAGAGC CTGTTCAAGA GAGAGTTGCG
43751 CAAGTCGCCG GTGTTGTGTA GGAAGATCCG GATTTTGTGC GATTTTCTTG
43801 AACTGTTTTA TTTGGAAATA CATCGGTTCTG TTATCAATCC GATTGGGATA
43851 GGAAGCTACG AAATGAGGGT TAAAGTCGCC AAGTATTGGA TCAACTTGCA
43901 TGGCAGGAAG AGGCAGAGGC GCTCGTCTTA GTACCATGCT CACCATGCGA
43951 TTAAAGCATT GCCAATAGCC TTGACTAGAT GGGCGGAGAG GCATGGGGGT
44001 AGCCACCCGT ACGGGAGGAG GAGGGGCCTC TGGTGGTGGA GTCGGCTTTT
44051 TATCAGCAGG TTGTTTAGGG ACTTTTGGGC TCGCTGGTGA AGGAGCTTTG
44101 GGGGGAGGCG GGGGTGTGTC CTCTGGGGGC GCGGTGCCCC GCTTGGGAAC
44151 ATCGGGTTTT TTGTCTTCAC CATCCTTAGG CGGTTGTTTG GCAATTTCTA
44201 TAGTTTTTGG CTCTGGTCCT TTGGGAAGAG TGGGAGGCGT TGGCAAGCCT
44251 TCTTTTCTGA CAACCCGATG ATGCTTGTA TAGTGAATCA GAAGAAGCAA
44301 ACCAAGAGTA ATGATATGGA GCAGAACGTA TCCTATGGTA CGTAGAATTC
44351 TAAGTAACAG AGGGTCTTTA GTATCAGTCG TTAAGTGGTA AAAATTATTC
44401 TTGTTATTGG GCGGGCAATG TGGTGGGGAA TTGGATACAT ACGTTCAAAA
44451 ATTGCTCGTT TTTAATCAA AATTATTCAA AGTTAAACT TTTTCGAGTT

44501 TGATGTATTT ATTTTTTTAT GTAAACTTTA ACACATGATA GAATTTAGCG
 44551 TATAGAGCGC AAACCTTTCAT GATAAAACAA ATAGGCCGTT TTTTATAGAGC
 44601 ATTTATTTTT ATAATGCCTT TATCTTTAAC AAGTTGTGAG TCTAAAATCG
 44651 ATCGAAATCG CATCTGGATT GTAGGTACGA ATGCTACATA TCCTCCTTTT
 44701 GAGTATGTGG ATGCTCAGGG GGAAGTTGTA GGTTCGATA TAGATTTGGC
 44751 AAAGGCAATT AGTGAAAAAC TTGGCAAGCA ATTGGAAGTT AGAGAATTCG
 44801 CTTTCGATGC TTTAATTTTA AATTTAAAAA AACATCGTAT CGATGCAATT
 44851 TTAGCAGGAA TGTCCATTAC TCCTTCGCGT CAGAAGGAAA TCGCCCTGCT
 44901 TCCCTATTAT GGCGATGAGG TTCAAGAGCT GATGGTGGTT TCTAAGCGGT
 44951 CTTTAGAGAC CCCTGTGCTT CCCCTAACAC AGTATTCTTC TGTTGCTGTT
 45001 CAGACAGGAA CGTTTCAGGA GCATTATCTT TTATCTCAGC CCGGAATTTC
 45051 TGTCCGTTCT TTTGATAGCA CCTTGAGGT GATTATGGAA GTTCGTTATG
 45101 GGAAATCTCC GGTTGCCGTT CTAGAACCCT CGGTAGGACG TGTCGTTCTT
 45151 AAAGACTTCC CTAATCTTGT TGCAACAAGA TTAGAGCTCC CTCCTGAATG
 45201 TTGGGTGTTG GGCTGTGGTC TCGGCGTAGC TAAAGATCGT CCTGAAGAAA
 45251 TACAAACGAT TCAACAAGCG ATTACAGATT TAAAGAGCGA AGGGGTGATT
 45301 CAATCTTTAA CCAAGAAATG GCAACTTTCT GAAGTTGCTT ACGAATAGAG
 45351 GGTATTCTTA TGGCAACCTC TGTTCTTGTA ACTTCATCTA CTTCTGTAGG
 45401 AGAGGCTAAC TCCTCCAACG AAAGATTTAC TGAACGAACA TCGCGAATGT
 45451 ATTACGCAGC TTTAGTCCTA GGGGCTTTGA GCTGTTTAAT TTTTATTGCT
 45501 ATGATTGTCA TTTTCCCACA GGTCGGATTG TGGGCTGTGG TCCTCGGGTT
 45551 TGCTCTTGGA TGTTTACTTT TAAGCTTAGC TATCGTTTTT GCTGTCTCCG
 45601 GTCTCGTTTT AGGCAAGACT TTAGAACCTA GTCGAGAAGC GACTCCTCCA
 45651 GAAATTGTTG CGCAAAAGGA GTGGACTACA CAACAAGATG TCTTAGGGAA
 45701 TGAGTATTGG CGTCCGAGT TGATTTCTTT GTTCTTACGA GGGGATCTCC
 45751 ACGAATCTCT GATTGTTGAT TCTAAGGATC GATCTTTAGA TATTGATCAG
 45801 AGTTTACAAA ATATATTGAA ACTTGAGCCC CTATCTACGA CACTTTCGCT

45851 GTTAAAGAAA GATTGTGTCC ACATCAATAT CATTTTACAT TTAGTGAGAC
 45901 AGTGGAACCTT ACTGGGAGTG GATCTTAGTC CTGAAGTCAC TGCGCACGCC
 45951 GAGGAACTTC TACTCTTTTT GATAGAAGAG CAGTATTACT CTCCTGATAT
 46001 TTTGAAATTG ATTCGCTACG GAGATGCTTT ACAAGCAACG TCTCCTTTGA
 46051 TGGATTGGGC AGATTCAGGT TCCTTTAGTG TAGACGCAGA CGGGGTATTT
 46101 AGCTGTCGCA GAGAAGAATG TTCTCCTGAG GATGCTTTGG CGCAATTCGA
 46151 TCTTCTTTTG GCGTTGGAAA ATCCCGACAG ACGCTTCTTA AAGGATTCTT
 46201 TTCTTACCTA CATTTGGTCG TCTTCATTTT TTGAGAAGTT TTTACATCGC
 46251 CATCTAGAGA GCTTGCAAAG AAAGCTCCCA GAGACAGCGA TCGATGTCGC
 46301 CCGCTATGAA GCACAAATAC AAACATTTCT CTCTCGCTAT TTTCAGAAGC
 46351 TCGATTTGAT AAACGCAATG TCCTTAGATT GGGGATATAA CTGTGCTGAG
 46401 GGAGAAAAAT GTTATGAGAG CGCAAATCAA AGATTAGACA ACCTATTTAT
 46451 TGCTTTTTCT TCTTCTGTC CTGCTATGAA GCGGCTCTTT GACAAATATG
 46501 GTTCTGTGGT ACGGGTAGAT CGTAGGCAGA TTCGTGAGCA GATTCTTTCG
 46551 AACACTGAAA TCTTAGAAAA TGAGTCAGGG TTCCTCTGCA GTTTGTATGA
 46601 ATATCCTTTA TCCTATTTGA TAGATTGGGC TGTTTTGCTA GACTGTGTTT
 46651 GCGGTACCGA AATCTCTCTA GAAGATCAGG CCGATTACAC CGTTTGTTTG
 46701 CAAGGCTTGG ATTCTATGTT ATCTCAATTT GCGAGTCGTT TACAGTCTGG
 46751 ACAAAAAGTA TTGAATCCTA GAGATGTTTT AAGTGAACAG GCTGCGGTTA
 46801 TGCTTGTTCA TGGCTTGGA GCACAGGGCG TGTCGTTTCA AGGATTGAAA
 46851 GCTTTGATGT ATTTGACAGC CGTTCCCCAA AGAATGTGGT TAGGAGCATT
 46901 GCCTTTATTT GAATCTTTTC CTGTCTTTAA TCGGATGAAA GAATTTCTTG
 46951 GGAATCTCT GGGAGACTAG GTGAATTTGT ATCAAAGAAG GAACAAGATT
 47001 GCATGTTAGG TTCTTTGCCA TGTTATCCTG GTGCTGGCAA TATTGAAGAA
 47051 TACAAAAATA GGTATTTCTA TTGTCAGTTA TGTGCTGAGG TCGTTAGTCC
 47101 CTATGTTGTT CCTGTTATTG TAGTTGATGT GCAAGGGGCT CCTCCTACAG
 47151 GTATCTTGCA GGTCTGCGT TGTAAGCAAC ATAAATTTCA AGGCCTACCC

47201 GTACATGGCC CCATTACTTC TTTATGGGCT TTGGAGCCCG TGGGTAAGGG
 47251 AGCTCCGCAG CTGGAGTCTG CAATGTACGA GCTCTGTTCT CAAGTAAGGA
 47301 ATTTTGACAT CTGCTCTATT GTGAGTTGGG TCTTTGGTGG GTTGTGTATT
 47351 TTTGCAGGTC TGATTGTCGG GGTAAATGGT GAAGCCCCTT TGATTGCGGG
 47401 ATTAAGTGCT TGGGTGATTC CCTGTATCAT TGGAGGGGTT GGTGCCATTT
 47451 TATGCTTGTT TGCGATCTTG ATGGCGTACT TGGGAAGAGG GAGAGTCCGT
 47501 GAGTGGCTCA ATCTTTCACA CGAATATATA ACGCAATGTC ATTGTCGTCA
 47551 GATACAGGCA CATTCTCAAA ACTATTCTGT GATCACAGAG TATCCTGCAA
 47601 CCTGTGCATT ATCTCAACCG ATTACAAAGT TACCTAATGG ATCACGCAGA
 47651 GATAACTAAG CGTGTTCGTC AGTTATTTCT CACATTTTCT CATGAATCTT
 47701 TTAGTGCGCT GCACGAGATC CCTCTCGAAA ATTTTAAAG ATAGATACTT
 47751 GGAAACTATG GTTTAAAAAG CTATAGAGGA TTCTAAATTG GGGTCTAGC
 47801 AACTTCTTGA CTTTAAGATC CAAAGTTAAG AGACTGACTA ATTATTTTGT
 47851 TTTGCTTGTT TTTCCAGATG AGCAATTGGT ATGGTAAGAG ATATTCAGAG
 47901 TGAATCTATA GGGAAATTAG TATTTTATAG CACAGGAAAT CCCGAAGGAA
 47951 TTCCCGTGCC GTTTTGCTCA TGTAGAGTGT GTCAAAACAC AGGGATTCAT
 48001 CGTTTACGAT CTTCGGTACT CATTCAATAT CAAAACAAGA CTCTAGTGAT
 48051 TGACGCAGGC CCTGATTTTC GTACGCAGAT GTTAGTTGCA GGGGTTTCCG
 48101 AGCTCGATGG GGTATTTCTG ACCCATCCCC ACTACGATCA TATCGGTGGT
 48151 ATTGATGATT TACGTGCGTG GTACATAGTC ACGCAGCGTT CGTTGCCTTT
 48201 GGTCTTTTCT GCAAGCACCT ATAGATTTTT AAACAAGGCT AAAGAGTATC
 48251 TCTTCGCCAC TCCGAATGTA GAGTCTTCAC TTCCCGCAGT TTTAGAGTTT
 48301 ACAATCTTGA ATGAGGACTG TGGGCAGGAG GAATTCAGG GCATTCCCTA
 48351 TACTTATGTT TCCTATTATC AAAAGTCGTG CCATGTAACG GGTTTTCGTT
 48401 TTGGAAATCT TGCTTATCTT ACAGATCTCT GTAGCTATGA TGCAAAAATT
 48451 TTCAGTACT TAGATAATGT AGAGACATTG ATCTTGTCTG CGGGTCCATC
 48501 GGAAACTCCT ATTCCTTTTC AGGGACACAA ATCTTCGCAT CTTACTGTAG

48551 AAGAAGCCAA AGCTTTTGCG AATCATGCAG GGATAAAGAA TTTAATTATT
 48601 ACACATATCA GCCACTGTTT AGAAGCAGAG CGTGACCAGC ATCCAGAGGT
 48651 CACATTTGCT TATGATGGCA TGGAGGTCCT TTGGACACTA TAGATACGCC
 48701 CGGGGAACAG GGTTCCTCAAT CTTTCGGAAA TTCGTTAGGG GCCAGGTTCTG
 48751 ACTTGCCTCG TAAGGAACAG GATCCCTCTC AAGCTTTAGC TGTGGCTTCC
 48801 TATCAAAATA AGACAGATTC TCAGGTCGTT GAAGAACATT TAGACGAGTT
 48851 GATCTCACTT GCGGATTCTT GTGGTATTTT TGTTTTAGAG ACCCGTTCTT
 48901 GGATTTTAAA AACACCCTCA GCTTCCACCT ATATCAATGT GGGGAAGTTG
 48951 GAGGAGATCG AAGAAATCTT GAAAGAGTTT CCCTCTATAG GGACTTTGAT
 49001 CATAGATGAG GAGATCACTC CATCCCAACA ACGGAATTTA GAGAAACGCC
 49051 TTGGCCTTGT CGTTTTGGAT AGGACGGAGT TAATTTTGGA AATCTTTTCC
 49101 AGCCGTGCCC TTAAGTGCAGA GGCAAATATC CAAGTCCAAC TTGCACAAGC
 49151 ACGTTATCTC CTTCTCGTC TTAAGAGACT TTGGGGGCAC CTATCTCGGC
 49201 AAAAATCTGG GGGAGGTAGC GGAGGCTTTG TTAAGGGGGA AGGAGAAAAA
 49251 CAGATCGAGC TAGACCGTAG AATGGTCCGT GAGCGTATCC ATAAGCTGTC
 49301 AGCACAGCTG AAAGCTGTGA TCAAACAGCG TCGGAACGC CGTAAAGTAA
 49351 AATCTCGACG AGGAATTCCT ACCTTTGCTT TGATAGGGTA TACAAATTCA
 49401 GGGAAGAGCA CCCTATTAAA TTTGCTGACG GCTGCTGATA CGTATGTTGA
 49451 AGACAAGCTA TTTGCAACTT TAGATCCCAA AACGCGCAA TGCCTACTTC
 49501 CAGGAGGCCG TCATGTCCTT CTTACTGATA CTGTAGGCTT CATTCGAAAA
 49551 CTTCTCATA CTTTGGTAGC AGCATTTAAA AGTACTTTAG AAGCAGCTTT
 49601 CCATGAAGAT GTTCTTCTGC ATGTTGTCGA TGCTTCGCAT CCTTTAGCTT
 49651 TAGAGCATGT ACAGACGACC TACGATCTCT TTCAAGAGTT GAAGATTGAA
 49701 AAGCCTAGGA TCATTACTGT GTTGAATAAG GTAGATCGGC TTCCTCAAGG
 49751 AAGTATCCCT ATGAAATTAC GTTTGCTCTC TCCTCTTCCT GTATTGATTT
 49801 CAGCAAAAAC TGGGGAGGGG ATCCAGAATC TTCTTAGTCT TATGACGGAA
 49851 ATCATTCAGG AGAAAAGTTT GCATGTGACT TTGAATTTTC CTTATACAGA

49901 ATATGGAAAA TTTACGGAAC TTTGCGATGC CGGGGTGTG GCCTCGTCAA
 49951 GGTATCAAGA AGATTTTTTA GTTGTGAAG CGTATCTTCC TAAGGAGCTG
 50001 CAAAAGAAAT TTCGTCCTTT TATTTCTTAT GTTTTCCCTG AAGATTGTGG
 50051 AGATGACGAG GGTAGAGGGC CCGTCTTGA GAGTTCTTTC GGGGATTAGG
 50101 TAGTTTTCTT CTAGGACATC GAATCTTTGT TAGTGAGAAA AAGAGTGATA
 50151 TTTTAAAATA GCCACTCATC GCTAAATCTA TTGAAGTCTC TAGAGGTATA
 50201 TGACGGTTGC GGAAGTCAAA GGAACATTTA AGCTGGTCTG TTTAGGCTGT
 50251 CGGGTGAATC AGTATGAGGT CCAAGCATAT CGCGACCAGT TGAATATCTT
 50301 AGGTTACCAA GAGGTCTTGG ATTCTGAAAT CCCTGCAGAT TTATGCATAA
 50351 TCAATACGTG TGCTGTCACA GCTTCTGCTG AGAGTTCGGG TCGTCATGCT
 50401 GTGCGTCAGT TATGTCGTCA GAACCCTACA GCACATATTG TTGTCACAGG
 50451 TTGTTTGGGG GAATCTGACA AAGAGTTTTT TGCTTCTTTG GATCGGCAAT
 50501 GCACACTTGT TTCCAATAAA GAAAAATCCC GACTTATAGA AAAAATTTTT
 50551 TCCTATGATA CGACCTTCCC TGAGTTCAAG ATCCATAGTT TTGAGGGAAA
 50601 GTCTCGAGCT TTTATTAAAG TTCAAGATGG CTGTAATTCT TTTTGCTCGT
 50651 ACTGCATTAT TCCTTATTTG CGGGGGCGTT CGGTTTCTCG TCCTGCTGAG
 50701 AAGATTTTAG CTGAAATCGC AGGGGTGTG GACCAAGGAT ATCGCGAAGT
 50751 TGTAATTGCA GGAATTAATG TTGGAGATTA TTGCGATGGA GAGCGTTCAT
 50801 TAGCCTCTTT GATTGAACAG GTGGACCGGA TTCCTGGAAT TGAGAGGATT
 50851 CGAATTTTCT CTATAGATCC TGATGATATC ACTGAAGATC TGACCCGTGC
 50901 CATCACCTCA TCGCGTCACA CTTGTCCTTC GTCACACCTT GTTCTTCAAT
 50951 CGGGGTCGAA TTCAATTTTA AAGAGAATGA ACCGGAAGTA TTCTCGCGGA
 51001 GATTTTTTAG ATTGTGTAGA GAAGTTCCGT GCTTCTGATC CTCGCTATGC
 51051 CTTTACTACA GATGTGATTG TCGGATTTCC TGGAGAGAGT GATCAAGATT
 51101 TTGAAGATAC TTTGAGAATT ATTGAAGATG TAGGCTTTAT TAAAGTGCAT
 51151 AGTTTCCCTT TCAGTGCTCG TCGTCGTAAT AAGGCATATA CTTTGTGATA
 51201 TCAGATTCCC AATCAGGTGA TCTATGAGAG GAAGAAGTAT CTTGCTGAGG

51251 TTGCTAAGAG GGTAGGCCAG AAAGAGATGA TGAAGCGTTT AGGAGAGACT
 51301 ACAGAGGTGC TTGTTGAGAA AGTAACGGGG CAGGTTGCTA CGGGTCACTC
 51351 TCCTTATTTT GAAAAGGTTT CTTTCCCTGT TGTAGGAACG GTAGCTATCA
 51401 AACTCTAGT TTCTGTGCGT CTTGATAGGG TAGAGGAAGA AGGGCTGATT
 51451 GGGGAGATTG TATGATAGAT ATAATGCAAC ATTTTAAGCC CTATACTATG
 51501 GTCCCAGGAC AAAAAGTCCC TATTCCTGGA TCTTTGTTAT ATGCTCAGGT
 51551 ATTTCTACC CTGTGGCGTC TATTTTCTTC GAAACACGAA ATCTTAAATG
 51601 AGCAGACCTT ACAGGTGCAA GGGCCTTTAA AACGCTTGC TGTTTTCCAA
 51651 GATTTACATC GTGGGGGGCT TGCAGTGA CTGAGCGCT ACAAGTATTA
 51701 TCTCCTCCC TCGGGAGAGT GCACACAATC TATCAAAGGG AACTGCCTT
 51751 CGGCAGCGCA AGCAGGGCCC CTGTTATCTC TTGGGGTGCA TAAGCATGCA
 51801 GATTGGCAA AGGTCCGTTG TCGTCGTGAT CTAAAGAAA TTCTTCCCCT
 51851 ATGGTCCGT TTCGCCGCTA TGGCTCCTAA GGGATCCTAT CGGGATCTAG
 51901 AGACGACGGC TATCGGTAGC TTGGTAAAGA CTGCCCATCA AAGAGTTTTA
 51951 CATAGGAAA CTACAGAGAT TGCTCCTGCG TACTCTCCA TAGCCCTTGC
 52001 GGGATTTTCA GAGTGCTTTC TTCCTAGGAG CTATGATGAA GAGTTCCAAG
 52051 GAATCCTCCC CCAAGATGGA GATCCAGAGG GGGGAGTTCC TTTTGAGCTT
 52101 CTCTCGTATA GCTTTGGTAT GATCCAAGAT ATTTTCTGA GACACCAGGG
 52151 ACAGCTAGTA GAGATCCTTC CTGCATTACC TCCTGAATTT CCTTGTGGCC
 52201 GCTTGATTCA TGTTGCCCTT CCTAATCTTG GACTTTGTC TATCGTCTGG
 52251 ACTAAGAAA CTATCCGTCA GGTGAGCTC CATGCAGAAT ATAGTGGCGA
 52301 GGTATTTTAA AAGTTTGTG CTTCACTATG CAGTGC GCGC CTTCGGGAAT
 52351 GGTGAGAGC ACGTCTCTCT GGATCTAAGA GACTTTCTTT AGGAGAACT
 52401 CTGGAGATAA AAGCAGGAAC CACATATTTA TGGGATTGTT TTCATAAATA
 52451 GATAGCCTTC CATGGTTGAT AACTGATCC ATCCTTGGGA TCTTGATCTG
 52501 CTCGTCTCAG GACGACAGAA AGATCCCAT AACTCTTAG GGATCCTTGC
 52551 TTCTGAAGAT TCTTCAGATC ATATTGTTAT TTTTCGTCCA GGGGCGCATA

52601 CGGTTGCTAT TGAACCTCTA GGAGAGCTTC ACCACGCTGT AGCTTATCGT
52651 TCGGGGCTCT TTTTCTTATC CGTTCCCAAA GGAATCGGAC ACGGGGATTA
52701 CCGTGTGTAT CATCAGAATG GACTTCTCGC TCATGATCCC TATGCGTTTC
52751 CTCCTCTGTG GGGAGAAATT GATTCTTTTT TATTCCATAG AGGAACGCAT
52801 TACCGCATTT ATGAACGCAT GGGGGCAATC CCTATGGAAG TTCAAGGAAT
52851 CTCAGGGGTG CTCTTTGTTC TTTGGGCTCC CCATGCGCAG AGAGTCTCTG
52901 TAGTCGGAGA TTTTAATTTT TGGCATGGCC TTGTCAATCC TCTACGTAAA
52951 ATTTCCGATC AGGGGATCTG GGAGCTTTTC GTCCCAGGCT TGGGAGAGGG
53001 AATACGGTAT AAGTGGGAAA TCGTTACCCA ATCGGGGAAT GTGATTGTAA
53051 AAACAGATCC TTATGGGAAG AGCTTTGATC CTCCACCCCA GGGTACAGCT
53101 CGTGTTGCGG ATTCTGAGAG CTA CTCTTGG AGTGATCATC GTTGGATGGA
53151 GAGGCGCTCG AAGCAGAGTG AAGGGCCCGT CACGATCTAT GAAGTGCACT
53201 TAGGCTCTTG GCAATGGCAG GAGGGAAGGC CCTTAAGCTA CAGCGAAATG
53251 GCGCATCGCC TTGCTAGCTA TTGCAAGGAA ATGCACTACA CTCATGTGGA
53301 GCTTCTTCCC ATTACGGAGC ATCCCCTGAA TGAATCTTGG GGCTATCAAG
53351 TGACGGGATA TTATGCTCCA ACATCAAGAT ACGGGACTCT CCAGGAGTTT
53401 CAGTATTTTG TAGACTATCT ACATAAAGAA AATATTGGTA TTATTTTAGA
53451 TTGGGTGCCG GGACATTTTC CCGTAGATGC GTTTGCTCTT GCCTCTTTTG
53501 ATGGGGAGCC TCTCTACGAG TACACGGGGC ATAGTCAGGC TCTTCATCCC
53551 CACTGGAATA CGTTTACCTT TGACTACAGT CGTCATGAAG TGACCAACTT
53601 TTTACTAGGG AGTGCTTTAT TTTGGCTCGA TAAGATGCAT ATTGATGGCT
53651 TACGTGTGGA TGCTGTGGCC TCTATGCTGT ATCGTGATTA TGGCCGTGAA
53701 GATGGAGAAT GGACGCCTAA CATCTATGGA GGTAAGGAGA ACTTAGAGTC
53751 TATAGAATTT TTGAAACACT TAAATTCTGT AATTCATAAG GAGTTCTCTG
53801 GAGTGCTCAC CTTTGCAGAG GAATCCACAG CGTTTCCAGG AGTCACTAAG
53851 GACGTAGATC AGGGAGGTCT GGGGTTTGAT TACAAATGGA ACTTAGGTTG
53901 GATGCACGAT ACCTTTCATT ACTTTATGAA GGATCCCATG TATCGTAAAT

53951 ACCATCAGAA AGATCTGACA TTTAGCCTTT GGTATGCCTT CCAAGAGTCT
54001 TTTATTCTTC CTCTCTCGCA TGACGAGGTG GTCCACGGTA AGGGCAGCTT
54051 AGTGAATAAG CTTCCCGGGG ATACCTGGAC CCGATTGCT CAAATGAGAG
54101 TGCTCTTGAG CTACCAGATC TGTTCGCTG GGAAAAAGTT ACTGTTTCATG
54151 GGTGGGGAAT TCGGACAATA CGGCGAGTGG TCTCCTGATC GTCCCTTAGA
54201 TTGGGAGCTT TTGAATCATC ACTACCACAA AACTTTGCGA AACTGTGTCT
54251 CTGCATTGAA TGCCTTGAT ATTACCAAC CCTATTTATG GATGCAAGAG
54301 AGCTCTCAAG AGTGCTTCCA TTGGGTAGAC TTCCATGATA TAGAAAACAA
54351 TGTCATTGCC TATTATAGAT TTGCAGGCAG CAATCGTTCT TCGGCGCTTC
54401 TCTGTGTCCA TCATTTTCAGT GCGAGTACTT TTCCTTCCTA TGTTTTAAGG
54451 TGTGAAGGTG TAAAGCATTG TGAACCTCTT CTCAACACTG ATGATGAGTC
54501 TTTTGGAGGC TCAGGGAAGG GAAATCGGGC TCCTGTGGTC TGTCAAGACC
54551 AAGGGGTCGC TTGGGGTTTG GATATAGAGC TCCCTCCTTT AGCTACTGTG
54601 ATCTATTTAG TTACTTTTTT CTAAAAATTT AAATACTTTA TTTGTAAATT
54651 GTTGTGGGAT TGTTCTATTT TGTGGTGTAG TTGATATTAA TAATTTATTT
54701 TATAATTAAA AATAATTATT AGTATTCTT TTATGTCTAC ATCACCAATT
54751 AGCAACGATC CCCGATATTT GTCTTTGTCT AATGCAACTG AGAAAACTTC
54801 TCTTCTTGCA AATAGCCGGA GTCTCTCGCC AGTACCAAAT TCCCTAGTTC
54851 CTAGCAATCC TGAAGATACA GGATTGCGAA AAAGTATTTT CACCCATTCC
54901 GTGACTTTAT TTGCTGGCCT GGTGTTTTTG CTGGTAGCGG TTTCTGTGTG
54951 TGTTGTCGCT TTGACCGTCT TAGCTCCCGG AGTTCCTCAG GCTATTCTTC
55001 TTGGAATCGC CATTTTCAGG GTGGGTATTG GTGGATTTTC TATAATGAAG
55051 AGCTTGTTTT ATATGGTCCG AGACTATATG TCCCCCAGGA TGCAGGAGTC
55101 GAGCAGAATC AAAAGTGCTT TAGCTGTAGG GACTGGATTT ACTGTCATGG
55151 GTTTGGTCAT GAAGGTGGGG GCGAATTTTG TTCCTGGAGG GTATGGGGGT
55201 CTCGTGGGTA GCTTGGGATC CAGTGCGTAT TCCCGGGGAA GCCAAACCAC
55251 ATTAGCAAGC TTCAGTCATT ATATTTATAC TAAGTTTTTC CGTTCTGAAA

55301 AAGTTGCTAA AGGGGAGAAG CTTACAGAAG CAGAAACTAT AAAAGAGGCG
 55351 AAAAAATTAC ACTATATCAC GTTGTCAATT GCCACTATTG GCGTTGGTCT
 55401 TGCGGTTTTG GGGATTCTCC TTGCCATTGC AGGAACGGTA TTGCTAGGAG
 55451 GCGCTCCCGC AACGATTGCT ATTATTTTAG CTCCCCCTTT AATTTCTATA
 55501 GGGCTTACGA CGGTTTGGCA AACGATACTC CATAGTAGTA TCGGAAAGTG
 55551 GAGAGCCTTT CTGCTTACTC AAGAAAAAAA AGATCTTTTT GTAGACACCT
 55601 CCCTGAAAGA CATTGCTTA GAAAAATTGC CCCCAGTGA GGTGGAAGAG
 55651 AGTGAAACTT CCCAATCTGT GATAGAAGTT CCAGATTCAG AGGGGATTGC
 55701 AGAGACGAGG ATCTCTGCGG AAGAAATCGA TACGAGGCTT TCCCTGACGA
 55751 CAAGACAGAA GGTCATCTTT GCTCTTGCGA CACTCTTGCT CTTAGCAAGT
 55801 ATTGCTGCCT TCATAGTCAC GGGATTGGT GGATTGACAG TCATGCAAGT
 55851 TCTCCTTGTT GCTTCTGTAG GATCGGCGGT TGCTTCTGTA ACACTCCCTA
 55901 TGGTTTCCTC AGGATTTTCC TACGTCGCCT ACCAACTGAA AGCAAGATTG
 55951 AATATCAGTA AATTACGTTG GAAAGAAGCA AAAAATAAAA AGCGGGTGCG
 56001 CCAGTTCTTA ATTGAGCTG GAGTGATTGC CTCGGATCGA GAATTTAACC
 56051 AAATGTGGAA GACAGTCTAC AAAAAACAGA TTCAGAAGAC TGACGCTGCA
 56101 ATTCGTGAAG AGGTTCGCAA TTTTGAGAAG GGTGGGGAAG TGAACAGCGC
 56151 CCTTGTTGGT GGAATCTTAC TTGGTGTAGG AACTGGGATC ATGCTTCTTG
 56201 CCCTGGTCCC TGCATTTGCT CCTATCGTTC CTGGTATTCT TGCTCTTGGA
 56251 GGATCGACGT TAGGAATCGC GGGATCGATT TTAATGAGGA AGTTTGTCAG
 56301 CTGGCTCTAT GATGAGCTTG TGAAGCTCTA TGAGCGTCGA CGTAATCGCC
 56351 GTGAGCTTCT CTATGGTCCT GAAAGTAAAA TGCGCTCCAT TGCTACGGAT
 56401 TTAGTTGTTG AGGCTCTTGC TGCTAGCCAC GATCATCTAT TTGATCTTGA
 56451 TGGTCCCGTA GATTTTATTG ATGTGGATGT AGATATAGAT GGAGCTGCTT
 56501 AGGCCAGGTC CTTGAATGTA AGATCCTCGA GCTTTGGGAG CTTGTCTGCC
 56551 TTCTCTAATT TTATTTTCTC TTTTACATCT AGTAGTTTAC TTAATTATAT
 56601 AATCAGCATT CTTTTTGTTT ATTTTAATTT ATATTTTGTT TTTAAAATAT

56651 TTTTATTTT AACATTTGTT TAATAGTTT TATTAAATAA TTTATCATTT
 56701 TAAGGTTCAA TTATGGCAGT TGGTGGCGTA GGCGGCTCAA GATCTCCTTC
 56751 CCCCATTCCCT CCTAATAGAA GGAATAGTGA GGATGGAAAA GTAAGTCCTA
 56801 AAGACAACCT AGGGGAACAT ACAGTTAGCA GTAGTGACAG TAGTCTTGCA
 56851 AGTCAGGGCC CTACAATAGA AGAGAGAAAA GCCCAGTTAG GCGGGACTGA
 56901 TAAAATTCCT TTGCCATCTG TCAAAGAACC CGGAGATTCT CAAACTTCAG
 56951 GACGTTCTGG GGTACTTCAG AGAATTTGGA AAGGCGTTAA AGGGGTCTTT
 57001 AAAAAAACC CTCAAGCGCG TCCTGAAGTT TCTAGTCCAC GTCTTCCATC
 57051 CCATGTGCAA CATGGCCAAC GTCTTCCTGG ACTCGAGGGC TTTAGAGATC
 57101 GTATCCAGAA AAGATCTGAA AATCCAGAGG CAGATTTAGG GAAGATGAAA
 57151 CGTTCCTATT CTGATGGTGA CCTTGATCGA GTAGGACACG ATTCTAATGA
 57201 AGATTCTACA GAGGATAGCC GTTCTGAAGG AGGAGAGCCT TCTTCAAAGA
 57251 GTTCTTCCTT CTTATCAGGA GTTCGAGGAG CGGTGTCTAA AGTTCATGGT
 57301 GCCCTAGGTG ATATTAAAGG AAAGTTCCAG CGTTCTGCTT CCGAAGATGA
 57351 TTAAACAAC CAGGGCGAAG ATTCTGCCGG CGATACTGTA AAAGAAAGGC
 57401 GTTCCGAAGA AGCAGAGGCT TCTTCGAAGA GTTCTTCTTT TTTATCAGGA
 57451 GTTCGAGGAG CGACGTCTAC AGTTCAGGGA GCCTTAGGTG ACGCTAAAGA
 57501 GAAGGTTTCG GCGTTCGGAG AGCAGGCTGC AGGTGCAATC AGATCAGCAC
 57551 CAGGGAATAT CAGAACTAGA TTCCAACGTT CTTCATCGGA AGGTGATCTT
 57601 TCTAATGTGA ATAAAGCAGC AAAACATCTG CGTAAGGCTT TAGAAAATTT
 57651 GGAAAAAGTA GCTCCAGAAC AAGTGTACC AGAGGTGGCT TCTAGGGTGC
 57701 AATCTCTTCT TGCACGCATG GAGCAATTGA CTCATCAGGA ACCTCCTACT
 57751 GTGGAGGATC TTATTACTTT CGTAGAATCC AATGTAGGTA GTGATTCTGT
 57801 GGAGTATGCA TCCATCGTAC CTCAAGATGG ATCGCAAGCC CCAGCAGAGA
 57851 CTGCGGAAGC TCCCGAAACA GGTGGGGTAG AGGGATCTGC AGCGCAGGGA
 57901 GCATGGAAAG CGTTACGGGA TTTTGTAGTT AGCATATTCC AAGCGGTAGC
 57951 GAGCTTCTTT AGGGCAATTG CTTCAAGATT AAGTTCAGCA CGACGTGAAT

58001 CAGCTGTAGA TGATCTTGCA TCAGAAAGTA ATACACAATG GTTTGTGGAG
58051 CAAGAGGGCG TTTCAAATCC ATCGGCTGCA CCTAGCTTAT CTTTTGCGGA
58101 AGAGATCGCT CGTAGAGCTG CAGAAATGAG TAACAGAAAT GCCCAGAGTC
58151 TTGAAAAATT GGAATCAGGC AATGTGACTG ATCCTGTCAT TCAACAAGGC
58201 TTAGGATTAG CTAGATCATT TGCTCCAGAG GGACAGTAGT CGTTATCTCA
58251 CTGTTTCTCT ATGCGCAAGG GAAACTTGAA GAGTTTAAAT TAAAACTCTT
58301 CAATATGTTG ATTATTTTAA TATATTTAAA AGCATTTTTG TTGTTTTTTA
58351 ATAAAATTAA ATTGTTTCAG AAAAAAGATT ATTCTTTTTA GGAAGTGTTT
58401 ATGGCATCAG GAATCGGAGG ATCTAGTGGA TTAGGAAAGA TTCCACCTAA
58451 AGATAATGGG GATAGAAGTC GATCGCCCTC TCCTAAGGGA GAACTTGCCA
58501 GCCACGAGAT TTCCCTGCCT CCTCAAGAAC ATGGAGAGGA AGGAGCTTCA
58551 GGATCTTCGC ATATACATAG CAGTTCCTCT TTTCTACCAG AAGATCAGGA
58601 GTCTCAGAGC TCTTCTTCGG CAGCTTCTAG CCCGGGATTT TTTTCTCGCG
58651 TACGTTCTGG GGTAGACAGG GCCTTAAAAT CATTTGGCAA CTTTTTTTCC
58701 GCAGAGTCTA CGAGTCAAGC GCGTGAAACG CGACAAGCTT TTGTTAGATT
58751 ATCAAAAACC ATCACC GCGG ATGAGAGACG GGATGTCGAT TCATCAAGTG
58801 CTGCTGCTAC AGAAGCCCGA GTGGCAGAGG ACGCGAGTGT TTCAGGCGAA
58851 AATCCTTCTC AGGGGGTTCC AGAAACCTCT TCTGGACCAG AACCTCAGCG
58901 TTTATTTTCT CTTCCTTCAG TAAAAAACA GAGCGGTTTG GGTCGGTTGG
58951 TACAGACAGT TCGCGATCGC ATAGTACTTC CTAGTGGGGC TCCACCTACA
59001 GACAGCGAGC CTTTAAGTCT CTACGAGCTA AACCTCCGTT TGAGTAGTTT
59051 ACGTCAGGAG CTCTCTGACA TACAAAGTAA TGATCAGTTG ACTCCAGAGG
59101 AAAAAGCAGA AGCCACAGTT ACCATACAAC AGCTGATCCA AATTACAGAA
59151 TTCCAATGCG GCTATATGGA GGCAACACAA TCTTCGGTAT CTCTAGCAGA
59201 AGCTCGTTTT AAGGGGGTAG AAAGTAGTGA TGAGATCAAT TCCCTCTGTT
59251 CAGAACTGAC AGATCCTGAG CTTCAAGAAC TCATGAGTGA TGGAGACTCT
59301 CTTCAAAACC TATTAGATGA GACTGCCGAC GATTTAGAAG CTGCTTTGTC

59351 CCATACTCGA TTGAGTTTTT CTTTAGACGA TAATCCAAC CCGATAGACA
 59401 ATAATCCAAC TCTGATTTCT CAAGAAGAGC CTATTTATGA GGAAATCGGA
 59451 GGAGCTGCAG ATCCTCAAAG AACTCGGGAA AACTGGTCTA CAAGATTATG
 59501 GAATCAGATT CGCGAGGCTC TGGTTTCTCT TTTAGGAATG ATTTTAAGCA
 59551 TTCTAGGGTC CATCTTGCAC AGGTTGCGTA TTGCTCGTCA TGCAGCTGCT
 59601 GAAGCAGTGG GTCGTTGTTG CACGTGCCGA GGAGAAGAGT GTACTTCTTC
 59651 TGAAGAGGAC TCGATGTCGG TGGGGTCTCC TTCAGAAATT GATGAAACTG
 59701 AAAGAACGGG CTCTCCGCAT GACGTTCCAC GCAGAAATGG AAGTCCACGT
 59751 GAAGATTCTC CATTGATGAA TGCCTTAGTA GGATGGGCAC ATAAGCACGG
 59801 TGCTAAAACC AAGGAGAGTT CAGAATCAAG TACCCCGGAA ATTTGATTT
 59851 CTGCTCCCAT AGTGAGAGGT TGGAGTCAAG ACAGTTCCGT CAGTTTTATT
 59901 GTTATGGAAG ATGATCATAT TTTCTATGAT GTTCCTCGTA GAAAAGATGG
 59951 AATCTATGAC GTTCCTAGTT CCCCTAGATG GAGTCCTGCG CGAGAGTTGG
 60001 AAGAGGATGT TTTTGGAGAT TATGAAGTTC CTATAACCTC TGCTGAACCA
 60051 TCTAAAGACA AGAACATCTA CATGACACCT AGATTAGCAA CTCCTGCTAT
 60101 CTATGATCTT CCTTCACGTC CAGGATCGTC TGAAGCTCA CGTTCTCCGT
 60151 CTTCAGATCG CGTACGAAGC AGCTCACCAA ATAGACGGGG TGTGCCTCTT
 60201 CCTCCAGTTC CTTACCTGC TATGAGTGAG GAGGGGAGCA TTTATGAGGA
 60251 TATGAGCGGT GCTTCAGGTG CAGGTGAAAG TGATTATGAA GATATGAGCC
 60301 GTTCCCCCTC TCCTAGAGGC GACTTGGATG AACCCATATA TGCTAATACT
 60351 CCTGAAGATA ATCCATTTAC TCAGAGAAAT ATAGATAGAA TTTTACAGGA
 60401 GAGGTCAGGC GGTGCTTCCG CTTCTCCTGT AGAGCCTATT TATGATGAGA
 60451 TCCCATGGAT TCATGGCAGG CCCCTGCTA CACTTCCAAG ACCCGAGAAT
 60501 ACATTGACTA ATGTTTCGCT TAGAGTGAGC CCAGGGTTTG GACCAGAAGT
 60551 AAGAGCCGCT TTGCTTAGCG AGAGCGTGAG TGCTGTTATG GTCGAAGCAG
 60601 AGAGTATTGT TCCTCCAACA GAGCCGGGGG ACGGAGAATC AGAATATCTA
 60651 GAGCCCTTAG GGGGACTTGT AGCTACAACG AAAATCTTAC TACAAAAAGG

60701 ATGGCCTCGT GGAGAGTCGA ATGCTTAGGA TTTAAGTAGT TCTTTCGAAT
 60751 CTCTAGTGAG GATGTATCGG GTTCTTAATT TTTATGGGGG AAACGTATCT
 60801 GTGGATTTC CTTAGCTTCT CCCATAAGAT TCATGATGGT AGAGAGTGTT
 60851 CACTACTCTC TATTTGGTCT TTACAGGTTG CATTGTCTAT ATAACATGCT
 60901 TTTAAACTT AAAGGTCGTT CCTGCGTGAA GGTAATGTGT CGTCGTTGAT
 60951 GAGGATACCG AGCCTTGATA GTCTAAGAAC ACCGAAAGTT TAGGGAAGAT
 61001 AAAAATTTGG TTTCTTCCTT TAAAAGCAAT GGCATTGCGA GCAAGGGTGG
 61051 TTCCTGATAG GAGCCATGAC GATCCACTAG ATTCTAGACT CACGTTGATC
 61101 TCAGGATTTT GTTGGTAGAG GACAGGCTGA TAAGCAAGCT CTATGTTCCA
 61151 ATAGGTAGGA AGACGGAAC TGGATTCCCA AGCGCTCTGA ATTCCCAGAG
 61201 GGACTGTCAG GTTATATAAG GGTTTATGAA CAGAAAATTT TCTAGCTTTA
 61251 TCTCCACTTT CTTGAAACGC AGTTTGATTA GAACGAACGG CAATTGCTTG
 61301 GATAAAAGGA GTGAAGTGGA GAGGTCGTGA TCGCCATTGT AGAGATAGAG
 61351 AGCAAGAGAG AGCCGCCCCCT AATGTCGTAC TATAACATTT GCCTTCCGTT
 61401 TGTATTTTTC CAGAATATCC AGATGCTTTG ATATGGTGGT TGCTGTAGCT
 61451 GTAGGCTAGA GATGCAGATG TAGAGAATCT CTCTGCAGC CAAGGATTAT
 61501 TGATCTGGAG CGCTACAGTT GTCGTATGCG AAGCCACGGA ATTGTCGGAG
 61551 TGGCTCTCGT AGAGATTACT GAAAAGTTGG GAGAAGTTTA CACCAAAGCT
 61601 ATGATTAGAA GCAGTGTTTG AGGTTGTTCC CAAAGAATAA CCCGTAGCTT
 61651 CCATATGGAA TCCTTTCGCA TCATTGTTGC TATTTTGATG CACGAAGAGT
 61701 CGAGTAGCTT CTCCAGAAGC TGTAGGTGCT ATTTGGCCTT GCTGTGTTTG
 61751 ATAACGTAGT GTCGCAAATA AGTTATGGAA AGATTGCCAG AAGGCAGATA
 61801 GGGCAATGTC TCCTTTGTTT TCTGGGTTTA CCTTATATCC TGTAGGTGTC
 61851 CAATCACCAT AAAGCTGGCG ATGTAAAGTA TTCACAGTAT CTTCAGAAGA
 61901 GGTATCAGAA GTTGTGATTG TTTCGATCCA GTAAGGGGAC CAAACGCCTT
 61951 GGTAGCCGTA GTGTTGAGTT GTATTTAGAC CCTCAGGGTA GAAATTATCC
 62001 GTATTAATAT GTTTAGCTGT GACGTCTAAG AGATACAGAA GAGGAACTTC

62051 TCGGATAGGT TGGGCAAGGT CTGCAGTATC ATAGGGATCT AGGTTCTCGT
62101 CATCCAGTAG GCTCAAAGGT CCTGAGAGAT TGATTATAGG GTTATTATCT
62151 TCGCTATAGG GTGCTGATGA ACCTGTGGGG CGAATCCATA GCTTGGGAGC
62201 AACTCTGTTG CCTAAGATAG AGGGAAGGTT AATTGCAAGA TTATTGATGT
62251 TAATTACAGA ACCCACACTA CTGCTACTTT GTTCTTCGTC TGTGTAGAA
62301 AACACAGCTC TACTGCCTAA CCGTAGAGTC CCACCAAATT GATCAAATTT
62351 ATAGACTTTC CACTCTGCTC GATCTTCAAG AGCGAGTGTG CCGTTGTACA
62401 GTCCAATGTG GTTTCTGAAA TGTGAAATGA AGTCATCACG AGAAGTCGAT
62451 GTATCCGGAA TATATGTTGA GGAGAACAAG ATAGTTCCGA GGTGTTCTGG
62501 ATTAGGATTA AATTTTGGGA TAGAGTTTGT TATAGTATAT CTTTGTAGTA
62551 TGGGATCATA GAAGGTAGCA GAATGACCTT GACTTGCTCC AACTGTTAAT
62601 GAGACATTAC GCGTGCAGTT TACAGAAACA TGATTGCTGA AAGTATCTTT
62651 GAAGTGTCTA TTATTATAAA AAATAATATC TCCCTGATCA GCAAATAAAG
62701 TGCATGCACC ATCTTGACGG AGCATGATAG CGCCGCCCCA AGTTCCCTGA
62751 TTGTTTGTGA AATAGACGGG ACCACTGTCT TGTATAGTTA GAGATTGTGT
62801 ACAGATAGCA CCTCCATCTC GTGCTGCAGT ATTATTATCG AAGGCTGCAA
62851 TTCCTGGGTT GTCTTTTATA GAACAACATA TGCAATAGAT AGCCCCCTCA
62901 GAGGAATGGT TAGCAGAGAT GTCCGCTTCC ATGGCAAAAT TATTGTTGAA
62951 GATCACAGAA CCGGTATTCT TTGTAAGAAT GCACTCTTGA TGTA CTCTTA
63001 TTGCACCACC CAGACCTGAT TGGTTATTCA AAAAATAGAT AGGCTGAGAA
63051 TTATTCTCAA TTCTACAAGC ATTAGCGAAC AACGCGCCCC CCGCTGTTCC
63101 GCCTGCAGCA TTATTAAAAA ACAGGCAAGG GCCAGTGTTG TCCTTAATGT
63151 TTATGATTGC AGCTTGATT GCTCCTCCTG AAGATTTTGC CTTGTTGTTA
63201 ATGAAGTATG CGGTTCCCTG ATTTTTTGAG ATTGTAACAT TTTTCGAACA
63251 TAAAACAGCT CCCCCTGTAC AAGTATCAGC GAAATTACTT GCATTAGGAA
63301 AGCTTAAATT CCCAGAGAAA ATGATGGAAC CATGATTCTC AGAAAGATCG
63351 AAATTACCAT TCACATACAT CGCACCAGCT CTTTAAATAG CAAAGCTATT

63401 TAGGAAAAGA ATTTGGTTTT TTGTATTCGT TATGGCAAGT GATTTGCAAG
63451 AGAGAGCACC GCCGTCTTGA GAGAAGTTTT CGAACCAGCT TTCTATGGAA
63501 TTCTGGTGAT CGAGGACAAT GTCTTGGTTA GTGTCATCCC TAACTCCAAA
63551 AAGTGTGCT CTATGAGAGT AGGGAGTCAT GTTAGTAAGA GTATCAATTA
63601 GAGGGAAGAG TGTTGTGAGT TGATTTGCTT GATTATCAAA ATAGTCAGAC
63651 AACGGAGTCG CATTAAGGAG TATTGTAGTT TTACCTAAAA TTAAGGCTCC
63701 AACAAAGAAG GAAGATTTAC TAAGGGATCT GTTATTTTGC ACTGTTTATT
63751 CCTTGATGTT TTCTTTGTGG TTAATATGTG CAATGACTCT CAGCTTTAAG
63801 GTATTGACCT ACAGTTGACG AGGAGACTTG AGCTGAATAA TCTAAGGATA
63851 AGGTTACTCT TGAGAAAAGT TGGGAAGTAT TTTTATTTT TGCAGCTACG
63901 GAATTATAGG AGACGGGAGT TGCTTGTGTT GTCCATGTTC CATTGCTGAT
63951 GAGTAGTGTC GTGAACATTT CTGGATTTTT TCTGTATAGG GTAGGTACGT
64001 AGGATATTTT CGTAGTCCAT AGCATGGGGA TATGATGTGA AGTTTTCCAT
64051 TCAGAACGGA AGCCTATGGG AGAGGAAAGA TCTGTAAGGG GATGTTTTGG
64101 ATGGAATTTT CTTATATGGT CTCCAGTTTC TTGGAACGAG GCCTGGGAAC
64151 AGCGCAGAGC AATGGCACTG ATAAAGGGCT GGAGTTCGAG AGTGCGGGTG
64201 ATTCTAGCTG GTAAGAATGT GCAGTCTAGA GAGGCTACCA AAGTGTGGTT
64251 ATTAAAGAAG GCTTTGGACG ACCCTTTTAA GATTTCTGTA TAGTGGCAAA
64301 GCATATGGTG ATCTCCGTAG CTATAACCTA GGGATAGCCC TGTAGAGATG
64351 AAGTCCCTGA AGAGGAGACT GTCGAAGCGG AGTCCTGCAA AGTAGTTGTG
64401 GGAGGAAGTC GTACTTGAG ATTGACGTTT TCTAGTTTTG GAGAACATTT
64451 GTGCGAATCC TAAAGAGAAA CTATGTCGTG CTGCAGTTTT TGCTGAGGTT
64501 GTTGCTGCAT AGCCCGTAGT ATGGTTTCGG AAGCCTTTGC GTCCCTCGCG
64551 ATTATGTTGG TTAATTAGAA GCCCGAGTCC TTGCAGAGAG GCTTCAAGGT
64601 CATGCTCTTT GAGGTTTTGT GGAGGTAAGA TGCGGATTCC TAACAGAGCG
64651 TTATAGGCAG ACTGCCATAA GGTATTAGCA ATAAATTCTC CGTGACGTTC
64701 CGGGTTAGGG CGGTATCCTA CAGGAGTCCA GTCTACGTAG AGCTGCCTGT

64751 GGTTTGTATT GGTCTGTTCC GGTACTGTAG AGCTTGTTGT AGTCGTAGTT
64801 TCCATCCAAT AGGGAGACCA GATTCCCTGA TATCCATAGT GCTCATCTAA
64851 GTTCATGGCT TCTACAATGA GATTGCAAGT ATCGATTTTT TTTGCAGTCA
64901 CATCGAGGAG GTAGAGGAGG GGGGATATCC TTTCGAGGTT CAGAGAGATC
64951 TAAGCTATCA TAGGGGTTTT CATTTTCATC GTTTAGAAAA GTCAAGGGTC
65001 CTGAGAGAGT GATAGTAGAA GAAGTGTCTT CAGAATAGGT GGATCCTGTT
65051 AATGTAGGAT AAATCCAGAA CTTTGGAGCT GAGGCTTCTG ATTGTA AAAAT
65101 AGAAGGAAGA TTGATCGCGA TTGCATTAAA ATTTATGGAG CTCCCCGGGC
65151 CTTTCGTCCT GATTAATGCT GCGTTTCCTA AACGTAGAAT GCGCCAGTT
65201 TGCGATAGGG TTTTGCAAGA AATAGCAGCC CGATCTTCAA TAGCGAGCAC
65251 ACCCCTTTCA AGTCGTGAAG AGTTAGAAAA TTTTGATAGG AAGTTCAATG
65301 GATTTGTTGC GTTAGAATCT ACATTGATTC CGGAAAACAA CACGGTGCCA
65351 AGGTGATGGG GTTCATAATT AAATACTATA GGATCTGTTG TCGTCTGATC
65401 GTGATCTATA GGATCATAAA AGAGAATTTT ATAACCCTGT CTTGCTCCTA
65451 GTTTTAAGTT AATCCCCGGA GCAGCATAGA GTGCATTCTT ATATCCGGGT
65501 TGAGGAGAAG AAGATGTGAT TGTATTATTG TTAAATAGAA TATCGCCGTA
65551 GTCTGCAGAG AGGAAGAAAT TTTGAGGAGT ACTTCCTATA CCAGAAAGAT
65601 TGATGAGAGC CCCTCCTGAA GTCGCAGAGT TATTAATAAA TGCTGTGCGA
65651 CCGTTATTTT GGAAGATGAA AGATCTCGTG TGTATAGCTC CGCCGCTAAG
65701 TGCTGCCGTT TTATTGTTAA AGATAAGACC TTTGGGATTG TTCTCAATGA
65751 CTAAGGAGGT ACACATGATA CCGCCACCAC CGGGATATAG TTTTCTGAT
65801 GCTGTGTTAA TTGAGGATGC GGAGTGATTG CTGATCTCTA TAATTTCTTT
65851 ATTTGAAGAG ATAATGACTC CTAGGGCAGA AAATATACCA CCCCTTGTC
65901 TGAAGCATTG TCATTGATTT GGATGTTTTG ATAATTCCTT TCTATATTTA
65951 CGGCTCTGCA GAAGATGCCT CCTCCAAAGC TGGAACTCTC TAATGTTTGA
66001 TTTTCTTGA TAACGATAGG ACCTAAGTTG TCTGTGATTG ATAACTCAC
66051 CCCAACATAG ATAGCTCCTC CTCTATTTTT AGAGACATTA TCTAAGAATT

66101 GTCCTTCTCC TAAGTTATTT GTAATATAGC TATCTAGTGC ATGTATAGCC
66151 CCACCAAATC CTGAAGTTGT AGTAGTAGCT AAGGAAGCCG CATTTGAAAT
66201 AAACGAGTAG TTCTGATTCT TAGAAATCCA GCATTCCCGA AACTGTAGA
66251 GAGCCCCTCC AGAACTGTGG GAGCTATTCC TTTCAAACT TAAGTTTCCT
66301 TTATTTTCAG AAATGAACAA ATTTTACAT GCAAGAATGC CTCCATCCTC
66351 GTAGTGGTAG TTATAATCTA TAACAGAATT GATAGAGTTT CCTGTAATTG
66401 TGATGTCCTG GGTATATCG TGACCGAATC CATTAAGAAG ATTAGAGGGA
66451 CGAAAGGAGG AAACCAAGGG CTCTGGAGTT ACATTCAGAC GGAAGCTTGG
66501 AGACGTGTGG AAAGCAGAGA TGTCTTTTCT AGACATCTGA CAAGAGGCGA
66551 GGTTAGGGAC TTCATTTCTT GATAAGGAAC AACATAGCGC AGTGGATAAA
66601 ATGCTGAGAC AAATGGGTCG CATAAACATC CAACTTTGA TATGAAAATA
66651 GAGTTTGTC TAAAAATCA TCAATTTGTT GTCTAGAAGT TTATAAATTT
66701 ATATTTAAAT ATGCTTATTA ACATGTTGAA TATTACAGAG TTAGTAACT
66751 CATATATTTG AATGATTTTC TATATTTAAA TTTTGAGGGA GTCCTCTACC
66801 GATTGGGCTA ACGTAGCTCA TCGGATAGTT GTTAATTCTA AAGATTTCAA
66851 TTTTATCGTT CTTTTATTTT AGAATTTTAA GGTACTTCCT GCTTGGAGAT
66901 GGTGCGTAGA TGTCGAGGAG GAGACCGATC CTTGGTAATC CAAGAATAGA
66951 TCGAGAGAAC GGAAGAGCGC AGTTTGATTG TGGACTTTGT ACCCTAAAGC
67001 ATTGCGAACA TAGTTATGGC CTAGGATATC CCAGGAACCT CCGCTCGCAA
67051 GTAGCGTGAC ACCGATTTGG GGATTTTGTT GATAGAGTAC CGGTTGGTAA
67101 GAAAGTTCTA GAGTCCATTC TGTAGGTACG TGAATTTTG ACTGCCATTT
67151 TCCTTGATT CCTAGAGGTA AGGTCAGATT ATAGAAAGGC TTTTGAGAGA
67201 CAACTTTTCG GGGATTGTCA CCAATCTCTT CGAACGCTGT TTGGTGAGAA
67251 CGTATTGCAA TTGCCTGAAC GAACGGGCTG AGGTGAAGAT AGGATTTCTG
67301 TTGCCAAGGG AAAGAACAGC CGATAGCTGC TGCTAATGTA TGGCTATAAC
67351 ACGTCCCTTC TGCCTGTTCT TGATGTGAGG GATGTAGGCT GTGGAGGTGA
67401 TGGTCCCCAT AGCCATACGC TAACACTGTG GATGTTGCAA AGGCCTCTTG

67451 GAACCACGGA AGCTCAACAT AAAGTGAAGA GACTGTATTG TGAGCCGAGA
67501 CGTTGTTGCT TGATCCGATT TCTTTAGTGC GGGTGAAGAA CTGTGCAAAA
67551 CCTAAGGAGA TTTTCTGATG TAAAGAAGTT TCGGAGGATG CTTGTAAGGA
67601 ATACCCTGTA GATTGGATAC GGAATCCTGG AGCCCCGGGG ATGCTATTTT
67651 GATGAACAAA GAGGCCGTCG GCAATCCCTT GAATTTCTAA GAAAGGCCTC
67701 TCGATATCAG AATCACCAGT TCGATTATAA CTTCTTAATA GAGAGAACAT
67751 AGTATGAAAG GATTGCCATA GGGGAGTCGT AGCAAGATCT CCTTGGTATT
67801 CAGGATTGAC CTTATATCCT AAGGGAGTCC AATTGGCATA CAGAGCTCTG
67851 TAGAGGGTGT TTGCCGTCTC TATAGAAGCG TTATTTGTTG TTGTTATCGT
67901 CTCTACCCAA TAAGGAGACC AGATGCCTTG ATAACCGTAA TGCTCAGTCG
67951 CATTTAAGCT TTCAGGATGA AAGTTATCGG TATTGATATG ACGTGCTGTT
68001 ACATCCGATA AAGAAAGAAG ATGAATGTTT TGTAAGGCT CAGAGAGATC
68051 TATACTGTCG TAGGGATCGC GGTTCCTCCTC ATTTAAGAGT GTCAGAGGAC
68101 CTGATAAAGT AATTGTAGGG TTATTGTCCT CTGTGAAAGG AGCACTAGAT
68151 TGTAGAGGAC GGATCCACAA GGTAGGAGCT TTTCCTTTTG CTAAGATCGA
68201 GGGGAGGTTA ATCGCAAGGT TATTAATGAT GACCTGGGAG CCTACACTAG
68251 TTGATGGAGT CTCAGAGTTG GCAGTTGTTG CAATACTCGC CGCATGCCCT
68301 AATTTAAGGA TACCTCCTTT TTGAGTGAAC TTATAGAATT GCCATCCCGC
68351 ACGATCCTCG ATAGAGAGGA CACCATTGCG AAGTTCAGAG GTATTTTTCG
68401 AGCTGCTAAT GAAATTATTT TCGTAGTCAG AAGCTTCTGG GATATAGGCT
68451 GAAGAAAATA AGATCGTTCC CTGATGGTTC GCATTGGGAT TAAAGATTAG
68501 AGGATTTGTA GTTGGATGTT GGTGTTCTAT AGGATCAAAA AAAGCAGTCG
68551 TATACCCCTT ATTAGCTCCA AGTTGTAAGT TGCTATTTGG TGTACAATGT
68601 ATGGCGTTGT ATCTACCAAA TGTGGTGAGG AAAACCTCAT TATTTTGAAA
68651 TGCGATATTT CCTTGTTCCG CGAAGAGTAG GCAGGTGCTG TCCTGTAGGA
68701 GCATAAGAGC ACCTCCCCAG TTTCTTGAT TGTGGTGAA ATATACGTGG
68751 CCACTATTTT TGATTGTCAA AAATTGTGTA CAGATAGCTC CGCCATCGCG

68801 AATGCAGTAG TTATTATTGA AAAGAATAGT TCCAGGGTTA TCGTCTATGG
68851 ATAGGTTTGT TGTATAAATC GCCCCTCCTG AACCATTTC TGAATTTATC
68901 GAACCAGATA ACGCTGTGTT GTTATTGAAA ATCACCGACC CGGAGTTATT
68951 TTTTATCGCA ACAGTAACGC TTGTTTGAAT GGCCCCGCCA TTGTTCCAC
69001 AGTTGTCTT AAAATAAATA GGACGCGTGT TATCAGAGAT CGTTGTATTT
69051 TCACTACGAA GCGCACCCCC TCCACTAGGG GCTGTATTGT TAAAAAGAG
69101 TAGAGGTGCC CTGTTGCTTT GGATGCGGCA GTGTCCATTG GTGGAGAGGG
69151 CTCCTCCCCA GTTGTGACG GAATTGTTGA CAAAGTAGAA AGTCCCTTGA
69201 TTTTGAGAAA TCGTGAAGTC TCCATTACAG GCAATCGCAC CCCCACGAGT
69251 TTCTCCTCCT GTECTCGCAT TGTTAAGACC TCGATTGCTG AAAAAATAA
69301 GGGGTCCTCT ATTCTTCGTG ATTGTGCAGG CTCCCTGGCA AGCAATCGCG
69351 CCTCCAGTCC CAATCGCGAG ATTTTACTG AAGAAGGCAT GGTCTTCAAC
69401 ATTTGATAAT AAGAAATTAT TACAGGACAC AGCTCCCCCA GCCGATGTCC
69451 AAAGAAGAAG GATGTTATCA ATAGACTTGT AGTTAGAAAG TACAATGTCT
69501 TGAGAGGAAT TATGTCTATT TCCAACAAAC GTAGTTATTG GAGAAAATCC
69551 TGTAAGAGTG GAGAGAGAGT CTAAGAGAGG AAAGCTCGTA CGAACTCTT
69601 CATCCCTCTC TAAAGCAAAC TTTTCAAGGG AGTCCGTTG TAACTATAC
69651 ACTGCAGGAG TCATCCCGAA CATGCAGGCT GTGAAATTCC CGAGATAGAA
69701 TAAAACTTA GGAGGAGTCT TTGACACTAG AGGTTCTTA ATCTTTCTGT
69751 TTTGGTCACT TATTGTAAA ATCTCATTCT ACTCGCCACG TTTAAGTAGT
69801 GACTCAGCGT GGAGGAAGAA ATATCCGCAG AGTAATCTAA GGAGAGAGTG
69851 ACTTTAGGAA ACACCTGCAT GGTATTTTTC ACTTTGATCC CTAAAGCATT
69901 GTAGGTCACA GGAGTGGCCT GCGTCGTCCA CGTACCTTGG CTAATCAGTA
69951 ATTTGAGTG GAGTTCAGGA TCTTGCTAT AGAGAGTAGA GCGATAGGAA
70001 ATTTCTGTGA GCCAGACTAG GGGAACTCGG TGGTGGTTCT TCCAAGAAGC
70051 GCGGATTCCT ACAGGGAGGG AGACGTCCGT TAGGGGGCGG TGAGGGAAA
70101 ATTCCCAGC ATGGTCTCCA GATTCTTGAA ACGCAGCAAG ATTCCTCGG

70151 ATGGCTAAGG CAGTAATAAA GGGATAGATC TGCAGGGACT CGCCGTGAGG
 70201 TTGAGGTAAG AAAACACAGG AGAGAGCCCC TGCTAAGGTA TGGTTGTGGA
 70251 AAGATCCCTG AGAGTTCCTT TCCAGGAGAC CCTGATACAT TGTATGGGTA
 70301 TGTTCGAGG TAAACATATA AGCAAGAGAC ACAGATAGAC GTATCCACTC
 70351 TTTGAAGAGA GTATTTTCTA TGCACATTCC AGAGAAATAG TGGTGAGAGG
 70401 ACGTGCTATT TTGAGATTCA TGTCTTTAG CTTTGGAGAA GAACTGAGCA
 70451 AATCCTAAAG AGAAATTCGG ACTTTGAGAA GAGGTTGCTT CGGTGGTAGC
 70501 ACTATAACCT GTCATATGAC TACGAAATCC CTTAAAACCG TTTTGTCTT
 70551 TTTGATGAAC CAGAAGACCA ATGCCTTGTA GGAAGCTGC ATGACCCTTC
 70601 TCTTCATCCC AGGAGGAGAG GGAGTGGAGT CCTGCAAGAG CCGTATATGC
 70651 CGATTGCCAC AAGGCATTCC TAATGAATTC TCCTCGACGT TCGGGATGAG
 70701 GACGGTAGCC TAGAGGAGAC CAGTTTGCAT AGAGCAGCTT GTGTTTTGTA
 70751 TTCGCGCCTA GTAGAGATGT AGGGTTCGTG ATTGTTGTAG TTTCTACCCA
 70801 ATAGGTCGAC CAGATGCCTT GATACCCATA GTGTCGCCA GAATTTAATG
 70851 TGGATAGATC CAGTTGCGAA GAGTTAATTT TTTGTGCAGC GACATCGACA
 70901 ATATAAAGAA GGGGAACTTT CTCAAGAGAG TGCGAGAGAT CCAGACTATC
 70951 GTAGGGATCT TCGTTGTTGC TGTTGCGTAA GGTGAGAGTT CCTGAGATTG
 71001 TGATTGTCGG GTTGAATCT TCAGTATAGG TAGATCCTGT TTTGTGGGG
 71051 TAAATCCAAA TTTTGGAGC CTGAGCTTGA AAAGAAAGAA TAGAAGGAAG
 71101 GTCAATGGCA ATGTGATTTA AAGTTATAGT ACTTCCTACT GTCGTTGGTG
 71151 TTGAGGATGG TGTGGGAATC GTTCCTGCTG TCGTGATCAC CGCACCTTGA
 71201 CCTAGAAGTA GAGTGCCTCC TCGTTGGAAG AACTTATAGC AGGCCAGCCC
 71251 CGCACCATCT TCAACAGCAA GGAATCCTTG ACGTAGTTCC GAAGTGTTC
 71301 TTAAATAGGA AAAGAAATTC ATTTTCATCGG TAAAGTTCTG GTGTACATGT
 71351 TCCCCTGAAA ATAAAACTGT ACCTGTATGA CCGGTTTCGA AATTAAAGAG
 71401 TATGGGGAAG GAGGAAGGGA GCTCATGTTT TATGGGATCA TAGAACAGCA
 71451 CTCGATAGCC GGGACGGGCT CCTATTTGCA GATTCATATT AGGAGTCGAG

71501 TGAATGGCGT TTCTGTATGG AGGATTGAGG GCATGCTTGG AGGCCGTATT
 71551 ATTGTTAAAG ATAATATCTC CATTATCTGC AGATAAGATG AAGCTTCCGT
 71601 TTCCAGAACC TGCTGATAAG TTGAGGAGAG CCCCTCCCCG AGTTGCAGTG
 71651 TTATTTAAAA AGTATACAGG ACCATTTTCT TTGATAATGA TAGATTTTCGC
 71701 ATGAATGGCT CCACCGTTGC TCTGGCTTTG GTTATTGTTA AAGAGTACCC
 71751 CTTCTCGGTT GTTCAATATC GTGCAAAAGG TGGTAGTAAG ACCTCCTCCT
 71801 CCTGGATTGA AGTTCGATCC ATAGTTATTT GCGAACGCGG AATTTTCACT
 71851 GATTTCTATG AGTTTTTTAT TCGAGGAGAT CGTGAGTGTT TGGGTAGAAA
 71901 ATATGCCTCC CCCAGATCCT GAAGAGTTGC TTGTGATCTG TATAGCTCCA
 71951 GAATTTCCCT CTATATTTAG AGAGTTCCCA CTATAAATCC CTCCTCCTAA
 72001 ACTGTCCGAA TTTAGTGCCC GATTCTGCTT GATTATGATC GGGCCTTTAT
 72051 TGTCTTTAAT AGATAAGTTC GTCTCAGTAT AGAGGGCACC CCCCTTATTT
 72101 AAAGCGAGAT TGTCAACGAA AGTTCCTGT CCTAGGTAT TAGTAATAGA
 72151 GCAATTTATG GCAAAGAGAG CTCCACCCAA TAGTGATCCC GCAGTGGCTG
 72201 TAGGATTGTC AGAGACCAAG TTTGTAGTAA ATGCATAGTT CTGATTCTTG
 72251 GAGATCGTGC AATTTTGAGC AGCATAAATT GCCCCGCCAG AATTGGGACA
 72301 GACATTCTTC TCAAAGAAGA CATTCCCTAT ATTTTCAGAG ATCAGAAGAT
 72351 TCTTACAGGT AAGAGCACCT CCATTCGACC GATAGTACTT ATAGTCCAAG
 72401 ATGAAATCAT TGTGATTCCC GACAATTGCG AGATCTTGAT TTTGGTTATG
 72451 AGTAAACCCT ACTTGAGGGG CTGCTTGATA TTCAGGACTT AATGTAATAT
 72501 AGGTCTCCAA AGGAAGTTGG AGACCTTCAT TAGCCAATAC AAAAGTAAAA
 72551 GGAAGCAACA TTCCGAAGCA AAAAAAGCGC ATACTCGTTC ACAAATAGAA
 72601 AGAAATATTC TGAATAAATC AATGCTAGAT TTTTGTCTA TAATTTGTTT
 72651 TATAGTGTCT AATTATGAGT AGTTTAAAGT TATCTTATTT AAGACAAGAC
 72701 GATTTTAATT TCCTTTTTTC ACAATAAAAT GCGAGTCCAA GAGTTTtagt
 72751 AAACGTACAT CATACTTTGG AAAATGGCGG CAAGTAGTCA GAGGATCAAG
 72801 CTCTTTGCAT ACTTTGCTTA GAGCAGTTTT TGGTTCTCAG GGTGTTTTGC

72851 CTTTAGGGAG CTTGTGATGT TTGGGATCGG AGAGCCTCCT TATATCCCTA
 72901 GGGGATGCCC TAGGAACCTCT TCCGAAACAC CGAGGGTCTG TTAGAGATAA
 72951 AAACAAAGGA CCATCGGGGA GACTTGTA CTATAAGAGCC ACTTATCTCT
 73001 ATTCCTATAG ATTTTGTCTG ATTTTGATTA TTAGAAAATA ATAGATTTTT
 73051 TCTGTTTTTT AAAGTAAAGT ATTTTATAA AGACTCATTT TTAATGAGTT
 73101 ATTACTTTTC TCTTTGGTAT CTGAAGGTGC AACAGCACTT TCAAGCAGCA
 73151 TTTGATTTTA CTCGCTCCCT GTGTTCACGA ATTTCTAATT TTGCTTTGGG
 73201 AGTGATTGCA TTGCTTCCTA TTATTGGGCA GTTGATATGA GGGCTGGACT
 73251 GGCTCCTCTC TAGGATAAAA AAGCCAGAAT TTCCTTCCGA TGTGGATCAG
 73301 ATCGTGCGAG TAGAACACGT CGTGGGTCAC GACCATAGAA GTCGAGTTGA
 73351 AGATATTCTA AAGAGACAAA GGCTCTCATT AGAGCCTAGA GACGAGGGGA
 73401 AGGTTACGCG AGATCTGCCT TCAGCTCCTT TTTTTTGATA TCCAAAGTCT
 73451 CAAGTTCCTA CAGTTGTTCT CTGAGGGGAC AGCTCTAAAT TTATTTCTGA
 73501 TATTTGCTCC ACTACGCAAC CGTGTGACTA CAGAATACAG TCGTGCTAGG
 73551 CAACCCGACC TACATAGAAT TGCCATCGTC TATATAGGAG TTCTCGATTC
 73601 AGAAAGTTCC AAGATCCTAG AGCGGCTAAT CTCTTATATG AGTTGTATCT
 73651 ATTCTGAATC GCAAATGTAT TTAAGATTCT TTATGGGCAA GAATGTAAAT
 73701 CAAAGTGCTG TACTCTCAAA ATTACATGTA GAAAATCTGC ACATCCGTTG
 73751 TGGGTTTTTC AGCGAGGATG CTGTTCCAGA GAGTGAGCCC TTCGATCTCT
 73801 CCATCTACGT GCACACAGAT CGTAGCTGTC CTCTCCCTAC GAAAAAACGG
 73851 AGCAGCTCCT GGGAACCTCA AACTGTAGAA CTCCCAGAGT CAATATATCC
 73901 ACAGTCGGAA TTCCTATTGA TGAGACCTCG AATGCTTTCTG TAGACTCTAT
 73951 GATGAAACAA GGAGTCGGGC AGGATGCTAA AGAGCTATAC ACATTTCTAT
 74001 CTCGTGGGAA TGAGCATTAC CAACCGTGTC TATGGTTCAG TCTCGAAGAG
 74051 GAACTCGGAT TCCTTTTCGA TGAAAAAATG CTCTGCGCCC CTCTATCTGA
 74101 GGATCACTAT TGCCACTCGT ATCTTGTA TAGTGGAT CAACATTTAA
 74151 AGGATTTAAT ATTATCGATG TTTTATAGAT CTCAGAATAT CTCAGCAGGA

74201 GAACTCCTCA AGGTCTCTAT AAACGTTGGA GATTCTTTTT CTCCTCTACA
 74251 ACAGAAAGAT TTCCTCTCGA TGGTCTTACG TGATGAAACG GGAAAAACG
 74301 TCGTCGTGGT TTTTAAAGGA GTTCTCTCCT TACCCGCAAC CCAAGTCTGC
 74351 AAATTAGTAG AGGAATTGAA CTCTAAGGAC TACTCCTACC TCAATATATT
 74401 TTCTTGTCAC GGAGATAGTA GTCCTCAGCT TTTATTCCGT AAGGAATTAG
 74451 AGGGAAGTTC AGGGCGTTAT TTTACAGTGA TTTGCGCTTT ATATCTAGGG
 74501 GATACAGACA TGCCTAGTTT ACAACTTGCT TCTGAAAGGA TCATGGTCTC
 74551 TAGAGAGTTT GATCTTGTA GATGCCTATGC TGCAAGATGC AAGCTCTTGA
 74601 AAATCGATCA TACAAATTGG AGACCTGGAA CTTTCAGTCG CCACGCCGAT
 74651 TTCGCAGATG CTGTAGACGT ATCAGCAGGA TTAACTCAA GAGAATTTAA
 74701 ACTGATTACG CAGGCGAATC AAGGGATCCT AGAGTCTGGA GAACTCCCGC
 74751 TCCCTTCAAA AACCTTCTGG GAAGGATTCT TAGCATTCTG TGATCGAGTG
 74801 ACTGTCACGA GAACTTTCAT TCCAATGTTA GACGCCGCTA TAAAGCAAGC
 74851 GGTATGGACT CATAAACATC CCAGCTTGAT AGATAAAGAG TGTGAAGCCC
 74901 TAGACTTGAA AACACAGTGC TTGCCATCTA TCGTATCGTA CCTTGAATAT
 74951 GTCACAACT CTCACGAAAA AACATCGAAA GGCCCGTTCA TACAAAAAGA
 75001 GATTATCGCA GACTGTTCTC CTCTTAAAGA GGCGCTCTTC CCAGGTTCTG
 75051 ATGAAGATGT TCCCTCTACC TCTGAGGATC CTTTCAAGTA TCATCCTTCG
 75101 GATCTTGAAG ACTCTTAATT AGTTGCGATA GAATTCAATT TTTTATATAA
 75151 AAATATCGT GTTGTCTTAA TTAAGAGATA GTTAATTTTC TATCTTTTTT
 75201 TAAATCTTTA TATAGCCTGC GTACGCTTTC ATTTTCAATG TTGGTTTGAT
 75251 CCTATGGCAT GCTATATTTT TATTTGGATA TCTACAGTTA AGCAGCATTT
 75301 TATTAGGGCT TTTGATTTTA CACGTCCTCT TGGTTCTCGG ATTACAAATT
 75351 TTGCTTTGGG GGTCATCAAG GCTATTCCCA TTTTAGGATG CGTTGTTATA
 75401 GGGGTAAGTT GGCTAGTTTC CACATGTTCT GCACGAAGGT TTGGGAAACC
 75451 GGCATTTACT TCTGACGTTG CTAGTATCGT GAAAATAGAA AAAACTCGAG
 75501 GTTATAATCC CCTTGCTTGG GTGGAACAGT ACTTGAGACA GCTTAGGGTT

75551 CGACTTCCTG AAGGAGATTT AGGAAAAATC CATGGGAAGG TCTCCAGAGA
75601 TTATGTTTGC GACAGGACTC CCCAAGAAAA TCTGAATATG GTTCCTCATC
75651 AATATCTGGG AGAGCTAGGT CGCGCGTTTT ATGGAATCCG CAACCGAGTA
75701 ACCAAGGCGT ATCAACGAGT CACTCCTCTG GAAGTCCCTT GTCTTACGCT
75751 CGTCGGTTTT GACATTTTAG ATCCCGAAGA TCAGGTGAAT TTCGTTTCGTC
75801 TGGCTAACGG CATACAAAC TACTACCCCC AAAC TCAAAT AAAACTTTAT
75851 TTAATCTCTA TCCAAAAGAT ATGGAATCAG TGTGACGGTA CGATTTCTCA
75901 AGAAAAAGAA CAGCAACTCC GCTCTCTAGG TTTGGATGCT AAAATCAAAT
75951 GTGTGTCGGC CCCCCTCTCT CTGCTCCAGA AATATCTTCA ATCCGAGAAC
76001 TTGCCTTCCT GTGATCTTCT CATTAATTAT TACGGGAAAC AACAGTCCGT
76051 CAGAGACGTG GACTCTATAA AGAGTCTACT CAATCTTTCT TCCGAACATA
76101 TCCCTGCGAT TTCTGTAACC TATAGACCTG ACGATCCTTT TTATAGCTAC
76151 TATTTCTTTC CTGGTTCTCA AGGAGGAACG GCACCCGATC AGAGGATCCC
76201 TTGGAGTGAG CAGGAGCATC TTCAAACGTA TACCACCCTG TCTAACCTTA
76251 GATGTGATAG ATATGCTGTT CACTTGGGAA TGGAAGATTT TGCCTCTGGA
76301 GTATTTTTAG ATCCTCTTAG GGTTCGGCT CCTTTATCTG GAGAGTATTC
76351 CTGCCCCTCA TACCTCTTAG ATTTAAAAAG TGAAGAGCTT CGTTGTTTCT
76401 TGTTATCCGC TTTTATAGAT CCCAACAATT CTGGTCAGGG AAATCCGCGT
76451 CCTATGTCCA TAACTTTTGG AAAC TCTCCT TTGGGTCAGA GGTGGTCTGA
76501 GTTTCTATCT CGTGTCTTAC ATGATGAAAC AGAAAAGCAT GTGGCTGTAG
76551 TCTGCAATAA TCCACAACCT ATAAAAAAGA GTTTTCCCTC ACATTCCTTA
76601 TCTCTATTAG AGAACGAACT GGAAGAGTCA GGTATTCTT ATTTGAATAT
76651 CGTTTCAGTG AGTCAGGAAC GCACGTGTGT TAAGGAACGT AGAATTTTAA
76701 GTTCTGATCC TTCGGGGAGG TCATTCACCTG TAATCCTCAC TGATCTTCCT
76751 GAAGGGAGTT CGGATATCCG CAACTTGCAG CTAGCGTCAG ATAGGATCTT
76801 AGTTTCTAGT GCTCTCGATG CTGCTGATGC CTGTGCTTCT GAATGTAAGA
76851 TCTTAGAATA TGAGGATCCC GAGCAAGAGT GGGCGCAACA GTATGCGTCG

76901 TTCTATAGAA ACATCGACAG GGCAGGCGAT CTTCAACGTC AGGGGATTCC
 76951 AGGAGAGCCT TTAGGGGTCT CAGCATCTAC GAGAGTAGTT TTAGAAAAGG
 77001 ACATCGTATT CAATCTCAAT GCGGTAATCC AACAGGCCAT GTGGAAGTTT
 77051 AAAAAACGGG ATCTTTTTGC TGTAGAAAGT CAGGCTTTAG GAGATGACAT
 77101 GCGACGTGCT TTAGAAGGTT ATATCGGCAG CAGTCTCTTA GTTGAGGGGA
 77151 CTATACAGCC TCAAGTCGCA TGTAATGTCA ATGTGAGTTT TGCTACGTTA
 77201 GACGAGGCTG TGTGTGCAGC TTGTGACTCA GCTCAAGATG CACCTTCTGA
 77251 GGAGAACAAAT ACAGATGACT AAAGATCGCA ATCTTGTGAA CGAAATCGCA
 77301 GATTGATGGG AACTAATTAG ACACACCTTT CTAAGGTGTT TGTTTTGATG
 77351 AACCTTTTTTA TTAGTCCAGC AGAGCTCTTT TTTGAAGATT CTTCTTTTTT
 77401 CTTAGGTCAT TCTGGGTTTT TTGAAGGTAT CGAGGGTTCT TATTGTCTAG
 77451 TTGTCTATAG AGGGTATCGA GGTTTTTTCT CTTAGGTATC CCACGATTCT
 77501 TTTGTATAGA AAAATTTTAT GAAAGCTTGA ACTCTTTACA CTGACTTTTT
 77551 ATTTTTCAAA TAAAAACGTT TTTAAAAATA TTATTATCAT AATTAGATAC
 77601 TTATTTGTTT TAATGTCTTA TTTGATTAAA ATAACCTTGT TAAAATTTTT
 77651 ATACATAAAT TTCTATTGTG GCTTGTCCTA GTATTTCTTC TTGGTTTACT
 77701 GTCGTTTCGAC AGCATTTTGT AAACGCCTTT GATTTCACCC ATCCCGTTTG
 77751 TTCTCGGATT ACAAATTTTG CTTTGGGGAT CATTAAAGGCA ATTCCCGTAT
 77801 TAGGACACAT TGTCATGGGA ATCGAGTGGT TGATTTCCTG GATTCCCAGA
 77851 CACACCGTTC GTCATGGAAT GTTTACTTCT GATGTCTCTA GTGCTATTAA
 77901 AGTAGAACAA ACACGGGGTC ATAATTGTTT AGCTCCCCTA GAAGCCTATT
 77951 TAAGTAGCTT GAGAGTCCCC ATTTCCCAAG AAGATCTAGG CAAAGTACAC
 78001 GGGAGAACCC CAGAAGATCC CTTCTAGAT ATCACACCCA CAGAAATTGT
 78051 CCAACTTCTC CCTGATGAAG AACTCTCTAC TGTAGATGAG GCACTGCAAG
 78101 GCGTTCGTAG TAGGTTAACC TATGCCTATA GGTCCGTAGA GAAACCTATG
 78151 ATTCAAGATC TTGCTCTTGT GGGTTTTGGT CTCCGAGATT CTGCGGACCT
 78201 CATAAATTC GTGCGTCTTG CTAATGGCGT GCAGAATCAC TATCCCCATA

78251 CTAAAGTGAA GCTCTATTTA GCGAAGAACT TGGCAGATGT CTGGGACTGT
 78301 GAAATTTCTG AAGAGGAAAA AGGGCAACTC CGAGCTCTAG GTTTAGACCC
 78351 TAAAATAGAG AGTATATCCC TTACGAGTGC AGGTCTTCCT TCAGTGCCAG
 78401 AAGTCGCTAC TGTCGATTTT ATGATTACCT GTTACGGGAA AGATCAGGAA
 78451 GTCCAAGATC CCTAGGTGAT ACAACATCTT CTAAACTTTG CTCTAGAAGA
 78501 GACCCCTTCC ATTTCCGTGC AATACCAAGA ACAAGAGAAG CTCTCTCCGT
 78551 GCGATCATTC CCCAGAAATA GGTAAGAAAGA AAAGATGGAA TAAGCTGGAA
 78601 TCCTTCTCCA CGTATTGTTC TCTGTTTATG TCTGTTAAGG ATCATTATAA
 78651 GCTGAATCTA GGAATTCAGA ATTCCCTGTC AGGGTGGCTT CTGGATCCCT
 78701 ATAGGGTTTG CGCGCCTTTA TCTTCACCGT ACTCGTGTCC TTCCTATCTT
 78751 TTAGATTTGC AAAACAAAGA GCTACGTCGT TCCCTTCTGT CAACGTTTCT
 78801 AGACCCTAAA AATCTCACTA GCGAAACATT CCGTTCTGTC TCTATAAACT
 78851 TTGGCAACTC TTCGTTTGA GAGAGATGGT CAGAGTTTCT ATCTCGTGTT
 78901 CTGCACGACG AGAAAGAAAA GCACGTAGCT GTTGTTTGTA ATGATGCAAA
 78951 ACTTCTGGAA GAAGGATTGT CCCCAGAGGC ATTGTCTCTA TTAGAAGAAG
 79001 ACTTAAGAGA ATCAGGGTAT TCGTATCTAA ACATTCTCTC GGTGAGCCCC
 79051 GAAGGAGTCT CCAAGGTCA GGAACGTCAG ATTCTAAGGC GAGATCTCCA
 79101 AGGACGGTCC TTTACTGTCA TGATTACAGA TCTTCCTTTA GGTAGCGAAG
 79151 ATATCCGTAG TTTACAATTA GCCTCGGATA GGATTTTAGT CTCCAGTTCT
 79201 CTTGATGCCG CGGATGCATG TGCTTCGGGA TGTAAAGTCT TAGTCTACGA
 79251 AAATCCAAAT GCATCCTGGG CTCAGGAATT GGAGAACTTC TACAAACAAG
 79301 TTGAGAGAAG AAGGTAGTGT TTCTTTCAGA GAATATTTCA GAGCCTATAT
 79351 GTGTGATAAA ATCGTGGCAC AGAAGAACTT CTTATTTACT TTAGACGCTG
 79401 TAATTAAACA GGCCGGTTGG AGATCACAAG AGAAACTCAA TTTATTTTAT
 79451 GTTGAAAGTC AGGCTTTAGG AAGAGAAATC AAAGTCAGCT TAGAGGAATA
 79501 TATTCAGAGT ATGGTCGGGA TTTTGGGATC TCAGAGAACC AAGAAAAGCT
 79551 TTAAGTTTTT TGTCGACTTT ACCCCTTTAG AGCAGGCTCT ACAAGAAAGA

79601 TGCTCTTCTG ATGATGACGA AGATGCAACA GCAACTTCGA CCGCTACAGG
 79651 GGCAACAGCA TCTCCGACTG ACATGCACGA AGATGAGTAA CGTTTGTCTG
 79701 ATACCTTAAA AGTTCCTTGC AAAGGGCTCC CTGAAAATA AATTCCCTCA
 79751 GAATCTCGAA TTCTCCTGAC TCTGAAACAA TCTTAGGTTT TCCTGAATAG
 79801 AATCTGACTG AAATTTCTGC TCGAATCTAA GGGCTGTTTC TTATTTTACC
 79851 CCTAGATGAG GATATTAAAT CCAAGCTAGG ACTTCAAAG TAGTTGGTTA
 79901 TTAGTTTATT AAAGAAAATA AATAAAAAA TATTTAAAAG CTGTTTATTC
 79951 AATTTAATTG ATATTTTCTA TGTGTATTAT TAAAATTGTT TGTTCCTAAT
 80001 TTTATTTTTT TTGTTGTTAT GCCAATTCCC TATATTTCTT CTTGGATTTC
 80051 TACCGTTCTGA CAGCATTTTG TTAAGGCGTT TGATTTCTCT CGTCCCTTTT
 80101 GTTCTAGGGT TACGAATTTT GCTTTAGGGG TCATCAAGGC CATCCCTATT
 80151 GTAGGACATA TTGTCATGGG GATGGAGTGG TTAGTTTCTT CCTGTGTTGC
 80201 CGGGATTATT ACTAGGTCCT CCTTTACCTC AGATGTCGTT CAGATTGTAA
 80251 AGACTGAGAA GGCCTTAGGT CGAGATCATA TATCTCGAGT GGCGGAGATA
 80301 TTGCAAAGAG AAAGGGGGAC CATAACTCCT GAGAATCAAG ATAAGGTGCA
 80351 TGGGAAGTTT CCTGTCTGTC CTTTGGTTCG TTTAAAATCC GAGGAACTT
 80401 TAAAACCTAA GCCGGGAGAA AGAGAGGGAA CTTTAGATAC TGTATTTTCT
 80451 CCGATTGCA CGCGCGTGAC TCGTGCGTAC TTACAGGCCC CCCGACCCGA
 80501 AATACGTACG ATTTCTATTG TGGGTTTCGA ACTTAAACT CCTCAAGATT
 80551 TCTCGCAATT TGTGAGTCTC GCGAATGAAA CGCAGAGACT GCATCCTGAA
 80601 GCGTTAGTTT GTCTGTATTT GACAGGCTTG AATCGCGAAT CTCAGATGTG
 80651 CGATACAACT ACTGCAGAGA AGAAGCAGTA CCTACATAAC TCAGGTCTCG
 80701 ACTCTAGAAT CCAAGTCAAA GACAGTAAAG AAGACGACGC TGGCTCTCCT
 80751 GAAAATCCCG AACTTTGGAT TGGCTATTAT TCACGAGAGC AACAGCATAA
 80801 TATAGACGGG CAGTATATTC AGCAGTGTCT AGGGAAGAGT GCAGATCCAA
 80851 TTCCTTGGAT TCATGTTACT GAAGACACAA AGGATTTTTA TTACCCACCA
 80901 AACTTTACTT CATACTCACA TACAAGACAA TCTACAGACC CAACATCGCC

80951 ACCAAGACTC CCTGAAAGTG AGGGGGATAA GGATTCCTTG TACGGACAAC
 81001 TGAGTCGATC GTATCACCAT GAGTATATGC TTGGTTTGGG ATTAAAACCA
 81051 GAGGATGCAG GACTCCTGAT GGACCCGGAT AGAATCTATG CTCCTCTATC
 81101 CCAAGGGCAT TATTGTCATT CCTACCTTGC GGATATAGAA AATGAGGATC
 81151 TACGAACTTT AGTCCTTTTCG CCTTTCCTAG ATCCTGGCAA TCTTAGTAGC
 81201 GAGGATCTTC GTCCTGTAGC ATTCAATATC GCTAGATTGC CATTAGAATT
 81251 GGA CTGTTA TTTTTCGCC TTGTTGCGGG TCAGCAAGAA GGGAGAAACA
 81301 TAGTTACCCT TGCCCACGGA ACTCCTCGTC CAGAAGATCT TGATCCTGAC
 81351 TCAATGAACA TTCTGACCAG AAGATTACAA ATGTCTGGAT ATAGCTATTT
 81401 GAACATTTTC TCCTATAAAT CACGGAAAAAT GATTGTAAAA GAACGTCAGT
 81451 TCTTTGGAGA TCGTTCTGAA GGGAAAGTCTT TCACATTGAT CTTATTTGAG
 81501 GATCCCATTA GTGCAGCAGA TTTCCGTTGT TTGCAGCTAG CTGCAGAAGG
 81551 TATGGTTGCT AAGGATCTCC CCAGCGTAGC AGATATTTGT GCCTCTGGAT
 81601 GTTCCTGCAT TCAGTTTTCT GAGATGCAGA GTCCTCAGGC TATTGAATAT
 81651 AGACAATGGG AGGCACGTGT CGAAGATGAA GCAGGAGAAG AAGCCAGAGA
 81701 ACCAGTAATT TATTCTCAGG ATCAATTGAG CAGCATGCTC ACTACACAAC
 81751 AGAATTTTGT ATTTTCTCTA GATGCTGTGG TAAAACAGGC GATCTGGAGA
 81801 TTCCGTTCTGA AAGGTCTTCT TACTATGGAA AGAAAGGCAC TAGGCGAGGA
 81851 GTTCTTAACT GCGATATTTT CCTATTTAGG GAGTCAGGAG CGTAATGAGA
 81901 ATATGGGGAA AAGAACTACC GAAGAACATG AGGTCGTTAT CAGCTTCGAA
 81951 GAGCTAGATC GCATGGTGCA AGTCCTCCCA GCCGAAGTCC CTGCAGATTC
 82001 AGGCAATGAT CCTACGCGTC CCGTTCCTAA TCCAGATAGT AACCCTGATT
 82051 CCTCGCAAAA TGAAGGCAGT TAGAAAGTAA AAATACTAGA GAAATTTCTT
 82101 ATCTCCAGAT GGAATCCGTG GTCCATGTAC CTAGGATTCC AGGAAGGGTT
 82151 TGCCTAGGAA TATCTTAATT TCACAAACCC CAGGAGTGCA CAAACTCCTT
 82201 AGAGGACTCT CTCTATATTT GTTCTTTCTA CACTAGTATT CCGTAGATTT
 82251 CTGATTTCCA GGATAGGATT CTAAATAGTT GTATCTAAGC GCTTTTACA

82301 ACCTTTCTCA GGTCTCCTCT TTCTAATTTT AAAAATAGCG AATTTCTTGT
82351 GGCTATAGGA GTCGCTAAGT ATCTTAGAGA TTGCTTTTTT CAAGTTTTTT
82401 TTAAATACTT CTTTCCTAAG TATTTCTTCC TAGTCGTGTT ATGGCTTCTT
82451 GTTTATCTGC CTGGTTTTCT ATAGTTCGTG AGCACTTTTA TCGAGCCTTT
82501 GATTTTTCTT TGCCGTTTTG TGCTCGTATT ACGGAATTG TATTAGGGGT
82551 CATCAAGGGG ATCCCTGTTG TGGGTCACAT TATTGTTGGG ATAGAGTGGC
82601 TCGTTTCTAG GTATTTAGAG AGTTTCGTGA CCAAGCCGAC ATTTGTCTCT
82651 GATGTGGTGA GTCTTCTGAA AACAGAGAAA GTTGCTGGTC GCGATCACAT
82701 TGCTCGTGTA GTGGAGACTT TGAAGAGGCA GAGAGTCGCT GTGGCTCCTG
82751 AAGATGAGGA TAAGGTCCAT GGAAGATTC CTGTGCATCC TTTCGGGGGA
82801 ATCCAACCTG TAGAAGTTCT CACTCTCTAT CCCGAAGTTC AAGATGCAAC
82851 GTTAGGGCTT GCCTTCTCTA AAATTCGTAA TCGTGTAAGA CAGGCGTATT
82901 TGCAAGCTCC ACGGCCAAAA CTGCAGAAGA TTTACATCAT AGGAAACGAT
82951 ATGAATCCTT TTGAAGTTGA CGACTTCTTG CATCTAGCCC GTCTCTGTAA
83001 TGAAACTCAA AGACTCTATC CTGACGCTAC GATTTCTCTA TATCTAACAG
83051 CTTCTGGTGG TCGCAATGCT ATGGACAAAA AGAATCGGAA GTTACTTAGT
83101 GATTGCGAAC TAAACCCCAA GATTGCTTGT TTGGACTTTA ATCAGGGTGA
83151 TGTAGTCAAA CAAGCAACTT GTGACTGTTG GATGGTGTAT CATGGGGAGA
83201 ATGATCAAGG TACGTTGAAT CAGATTCAGG AAGAGTTAGA AAAGTCAGGG
83251 GAGGAAACCC CTTGGATTCA TGTGGGGCAA AAGCCTCTTT CACAATCCTT
83301 GTGGGATTTT TCTCCATTTT CATCTTTGGA GATGAAGGGA GATAAAGAGA
83351 AAGCTCTAGA GACTCTGAA TTAGAAAAAG AACAGCTATA TTCTCGATTG
83401 GTATACGTAG GAGAGCGCTC TTCGGTCTT AGTTTGGGGT TTGGAGATAG
83451 TCGGTCAGGG ATCTTGATGG ACCCAAACG GGTGCATGCT CCCTTATCTG
83501 AAGGGCATTG TTGTCATTCC TACCTTGCAG ACTTAGAAAA TCCCGGGTTA
83551 CAAAAACAA TTTTAGCGGC ATTTCTGAAT CCTAAGGAGT TGAGCAGTAC
83601 CATACTGCAA CCTATATCTC TAAATCTTAT CTAAATAGC AAAACTTACT

83651 TAAGGCAGCA CTTTGGCTTT TTTGAGAGGA TGAGCAGAAG TGATCGCAAT
 83701 GTGGTTGTGCG TTGTATGTGA TTCTTGGTGG GGTACCGACT GGAAGGAGGA
 83751 GCCAAGCTTC CAACACTTTA TTATGGAGCT AGAGTGTCGA GGGTATTCGC
 83801 ACTTCAATAT TTTTGCCTTT AGATCTAATA GCATGTGTGT AGAAGAACGT
 83851 AGGATCTTAA ATGAAAGTTC TCAAGAGAAA GCCTTTACCA TGATTTTCTG
 83901 TGAGGATTCA GTATCTCAAG GAGATATCCG CTGTTTGCAT TTGGCGTCTG
 83951 AAGGAATGCT TTGTGGTAAA GAGTGCTATG CTGTCGATGT CTATACGTCA
 84001 GGATGCGCGA ACTTTATGAT GGAAGAAGTC TTAAC TTTGG AGCGAGAATC
 84051 TAATCTGTGG AATAGAAAGC ATGGTCTTTG GAAAAGAGAA GTTAGAAAAC
 84101 AGAAACAAGA AGCTGCTTTG GATCAAGACG AGAGCGAGAT TTACGTTTGT
 84151 AATCAGCTGA CGGCGCAACA GAACTTCGCT TGTTCCTGAG ATGCTGCAAT
 84201 CCGCCAGTCT ATATGGAGAT CCCGTATGCC AGAACTTCTC TCTATTGAGA
 84251 GACGGGCGTT AGGGGAACAA CTCTTTACTA CTGTACATCA CTACCTAACA
 84301 ACGCAAAAAA AGATCCTCAG GGAATCTAG AAACGCAGCA ATCCGCGCAA
 84351 TTGTCTATAG ATTTACAGC ATTAGATGAA GCTGTTGAAT CTCTAGGATC
 84401 GACTCTTAGC AGAGCTCCTT CAGAAATATC TCCAATTCCA GAGGAGGAAG
 84451 CTCACTTAGG AGCCAACAAA TAGAGACAAA GAAAATTCTGA CGGTTTGAGG
 84501 ATAACGATAC GCAATGTCTA AGCTTTGAAT CAGGATACTC TGCTTTACAG
 84551 GCGTGTCTTT GTGCCTATGG TCTCTCTCTC ATAACAGAGT CTCTCAAATC
 84601 TAATTGTCAG AACCTATTTT CCCTAAGAAT CGATAACGTA TTTGTTAGGA
 84651 AAACGTGCTC AACAAGAGCG TTTTATTTTC AGTGTACTTG ATGTCTAATA
 84701 CAAATTGTTT ATTTTGTATT TTTGGCACAT CATTTAAATT CCTTGTAATT
 84751 TTGAATCTAA ACGTAAATTT CTTATGACTC ATTGCTTACA TGTTTGGTTT
 84801 TCTGTAGTTC GTCATCACTT TGTGCAGGCG TTTAATTTCT CACGTCCTTT
 84851 ATATTCTCGA ATTACCCACT TCGCTTTAGG GGTGATTAAG GCCATCCCCA
 84901 TTGTAGGGCA TCTTGTTATG GGAGTCGATT GGTGATCTC TCATTGCTTC
 84951 GAGAGGGGAG TCTCACACCC TGGGTTCCCT TCAGATATTG CTCCTATACT

85001 GAAAGTAGAA AAGATCGCGG GCCGAGATCA TATTTCTAGA ATCGAAAATC
 85051 AGCTAAAGAG CCTTAGGAAA ACTATCGAGG TTGAAGATCT AGATAAAGTC
 85101 CACGGGCAAT ATCAAGAGAA TCCTTATGCA GATATGGCCT CTAGTGAGGT
 85151 TCTTAAACTC GATAAGGGAG TTCATGTTAG CGAGCTTGGC AAAGCCTTTT
 85201 CTAGAGTTCTG CAATCGCATC ACCAGATCCT ATAGTTATGC CCCTACTCCT
 85251 CAGTTGGACT CTATAGCTAT TGTGTTGATA GATCTCGTCA GTCCTGAAGA
 85301 ACAAGAGAAT TTAGTACGCT TGGCGAATGA GGTCATTCAA CTCTATCCCA
 85351 AATCAAAGAC AACTCTATAT CTTCTTATCG ATTTTAATAA GGAGTGGGTA
 85401 GGGGATATCT CCTCTGATAA GGAAAAACAG CTCCGTTCTC TAGGTCTACA
 85451 TTCTGAAGTT CAGTGTCTTT CCGTCTTGA ACCTCAGGGT GCCGAGGGCG
 85501 AAGATACGAA ACACTTTGAC CTTATGGTCG GCTGTTATGG GAAGGATTCT
 85551 TACTTAAGGG AGGGTAAAT TTTACAGCAG GCCCTAGGGA CTCGTTAGG
 85601 TACTGTTCCC TGGGTGAATG TTATGCACAC ATTGCCATCT AGGTATAGAT
 85651 CTCGGCTTTC CTTACCTATA AATACCGAAA AGGATAAGAC AGAGCTTTAT
 85701 AAAGAGATTT CTCGTACACA CCATCAGTTG CATACTTTGG GAATGGGACT
 85751 TGGAGCCCAG GATTTCAGGAT TGCTCTTAGA CCGGCAACGA CTCCATGCTC
 85801 CTTTATCTCA AGGGTCTCAC TGCCATTCTT ATCTTGCAGA TCTCACCCT
 85851 GAAGAGCTGA AAATTTTGTT ATTTTCAGCA TTTGTGGATG CTAAGAACAT
 85901 AAGTAAGAAA GAGCTTCGTG AGGTATCTCT AAATTTTGCT AACGATACTT
 85951 CCGTAGAGTG TGGCTGCGCT TTTTACTTTT AGTGTCTTAT GATGAGAAGG
 86001 AGAAAGACGT AGTTGTCGTT TGTAATCATT CTGAACCTAA TATCCTCGGC
 86051 CTGCCTCCTG AAGCAGTCTC TCAGCTTATT GAAGAGCTTA GCGATGAAGG
 86101 CTATAGCTAT CTGAATGTAG TGC GTTGTGA TCTCTCCGGG GAGACTACGG
 86151 TTCAACAACG TCTGCTATTG AATGCCGATG AAGGGAGATC TATGACGGTG
 86201 GTGATCTCAG AGCTTCCTGA AGGGCACCCC GATATTCGGA ATTTGCAGTT
 86251 GGCATCCGAA AGAATTTTGT TTTCTCGTGA AAAAGAAGCT GCTGATGCCT
 86301 ATGCTTCAGG ATGTAAAGTG GTCGCTTTCTG ATGATGAGCA TCTCCCTTGG

86351 GTCTCCAGTC ATATTGCCTA CGCGGAGGAG ATCAGAGAGA AACAGAACA
 86401 AACAAATGCAA GGGTCTTTAA CTGAAGAGCA GTTAGGAGCA CTCCTCTGCA
 86451 ACACAGTCTC CACAGAGAAA AATCTAGCCT TTGCTCTAGA CGCCGTGATA
 86501 AAACAGTCTG TGTGGAGATT CCGCAATCCG GATCTTTTTG CTTATGAGAG
 86551 AGAAGCTCTA GAGGCTTCAG TAACAGATGC TTTAGTATCT TACGTTTCAA
 86601 ATTTAGACAT GATACCGTAC ACAAGTTCTC AGGGCATAGT CATAGAAGAT
 86651 AGTAGTATCG TCCGTACCTC TCAAGAGCAT AACTCATTG TGAAGTGTGC
 86701 AGCATTCGAT AAGTTAGCGA GCCAAATAGA GTTCTTATGC CCCAGTGACG
 86751 TGTGCCCCAT TTCTGGTAAA GACCCTTTGA TTTCTGATGA TGAGGATGAG
 86801 GAACTGAATC CTAAAGTTTC ATCTGCTGCA GACTCTAAAG ATAAAACCTA
 86851 GGGAGTGAAT TCTACACGAG AATCGAGAGG AGAGCGAGTC TTTCAAGGAT
 86901 TCATAATCCT TGTTAACGTA TGCATAAACA AGTGAAGCCA TTGCAACGTG
 86951 AAGTAATCGC ATTGATGAAA GATGCTTTCC CTAGGGATGC AAAGCAGATC
 87001 GTATTCCCTT CTTTCCAAAA CAGACTATAG ATTCAAAGA TATTCTTTCT
 87051 TTTCAAATAG ACTTGAGAGA GGGGGGGGGG TTCTCATAGA GTGAGAATCT
 87101 TGGCCTTCAT TGCTAAGTTC TTCGATGATG GATAGAGGAT TCTAAGACGA
 87151 CCGGGGCTAC AGAAAACCTCT AAGCAGAGCT TAGAGTTTTA AAATGTGGAT
 87201 TTTAGTCTCG TAGACACTCG GTGGTTTGTA AATCCATTTT TCCCGTCAAA
 87251 GGTATAGTTT AGAAAGGCCT GAGTGTCTC GGTGAGATCT ACTACATCAT
 87301 GGATTTGTAC AAACAAACCG TGGCGGCGTA GATTTGCTCC TGAGATCGAA
 87351 GTGCTCTCTT GGTTTGAGAC GACAGTCACA ATATTGTGAG GGTTGACACG
 87401 ATAGATATCA GGCTTGTAGG CAAGTTTGAT GGTCAGTGTA GAAGGAGCCT
 87451 TCTTAAATGG TGTGAACCAT TGAGAAGAAC ATCCTATCGG TAGGGAACA
 87501 TTGTACCCTT TACCTCTACT AAAGCTACGT TGCAGATCTC CAGTTTCTGT
 87551 GAACTTACTT TGCCAACCAC CTAGGAATTC TGCGGAAATA AAACCTGAAA
 87601 GATCCCAAGC TTGAGCCAGA GGTCTTGTA GAAGACACCA GTTTAGGAAA
 87651 GGATGTTCTG CAGAAATAAG AACATAGTAA CTATTGTTAT GCCATTGCCC

87701 TTGAGATTTT GGAGCTTTGT CAGGTCTGAG GTAGGTGGTA TTTAGGTGAT
 87751 TTTTGGAAATA ACCATAAGCT GCTTTCCAGG AAATTAAGGC CTCGCTCTTT
 87801 TGAGTCACGA TAGGGAATTG ACCAAAGAAC GAGAGTAAAT ACATTTGTTC
 87851 TGAGCAACGT GAATCGTAGG GGTGGCGTT AGTTTTTCCA TAAAGCTGCC
 87901 CGAAAGAAAG TCCTAACGTA GTGTGGTCCG TGTAGTTCAT AGATAGCGCA
 87951 GCTTGGTAGC CTCCATAGCG ACCTGAAAAG CCCTCATGTC CTTGTCTTGG
 88001 TGTGTGTTTCG ACATAGGCTC CTAAAGCTTT CGCGGTTATG GACAACCCGG
 88051 GATGATCTAT CAAAAGAACA TCTTGGAGAA TATCAGAGAA TGCCTGATTT
 88101 CCTAAGAAGG AAATCCATAA GCTGTTGCTG ACAATTTCTC CGTAACGCTC
 88151 GGGATCTAAG ATATAGGTAG AACGCACGAG AGTGTCTGAA TTCCATACAG
 88201 CATAGAGAGT ATTTGCGCTA GGAGAGGGAC CTCCAGGAAA TCCTCCATCA
 88251 GGAGCTGGAA TTAACAGGGG ACGGGACCAT GTGTAGGACC ACTTTCCTTG
 88301 GTAGCCGTAG TGGCTTGGAG TCGCAATCTC CCCATCAGGA AATCCTGTCT
 88351 TAGTAACGGT TGCTCCTTTG AAAACAGCGA TAGGAATTGC TACTGGAGTT
 88401 TGTAATGACA CCATATCATA AAGATCTGTA ACGTCATGTT CATCAAGAAC
 88451 CAGAGCTCCT GTTAAAGTGA CGTTTTTTGT GCCTGCATTT ACTGATGCTG
 88501 AAACAAAATC TCTTTTTAGG AAGGAAAAAG GATCGAATGC TAACTTTCCA
 88551 ATCGTAAAGT CTACAGCGGC AGGTGCTCCC GTGGGTGTTG CCAGCCCTAA
 88601 GGTTCCTCCA GAGCCCAGGG TAAGCTGACC TGAGCCCTGA GTAGCGAAGC
 88651 CAAGAACATT GACAACCGCA TTGTCAGTAA TCTTCAGTTC TCCACTAGCG
 88701 ATCTTGACTG TTCCTAGAAG TATAGTTGTC GTGTTGGCAG GCAACAGGAG
 88751 TTCTGTAGAG GAGAGTCCCT TACTTGTAAG GACTACAGAT CCTGAAGCGC
 88801 CATTAGCGTT GATTGTAATG TCTTTATTAG ACGGACTTGT GGTGGGAGG
 88851 CTATGTGTAA TGGGATCATA AAATACAAGA CGTGAGCCTC CTTGTGCAGA
 88901 TAGAGACACA ATCTCTCCCC CTGCTTCTAC AGTGATGGCA TTGCGGATTC
 88951 CAGGTTTTGT ATTGAGCATG TTTCTTGGG ACGCAATATC ACCTTCGGAA
 89001 GCCAAAATCG TAAGTGTCGA TGTTTGCTTC GCAGGGTCTC CAACAGAGGG

89051 GCCTATGAAA ATAGCTCCAC CCTTCTCCGC GCGGTTTCTA GAGAATTCAA
 89101 TAGGACCGCG TCCTTGAATA TTGAGCACTT TAGCACAGAT GGCTCCGCCA
 89151 TTTCTCGCTG CAGTGTGCT ATCTAGGAGC AAGGCACCCT GGTTCCTAC
 89201 GATGTTGCAG GTTTCGGCGT AGATCGCACC CGCATCATTT GGTGTACCAT
 89251 TATAAGAGAA GGTGCACTTC CCCTGATTGT TTTTAAATTC GAAAGTTCCC
 89301 GTAGGGATGC AGATGGCTCC GCCACTTCCT AAAGCATAGG TTCCTGATGA
 89351 AGGGGTGACG CCTTTTAAGC TGTTACACA GGAGTTGGCG GTGAAGATGA
 89401 TGCATCCAGA ATTGTTTTCA AAAAGGAGAG AGCTTCCTCC ATAAATCACG
 89451 CCGCCTCCTG ACGAAGTAAA GTTATGGGAG AAGCTCATGG TAGCGGTGTT
 89501 ATTGATGAAT TTAACGACTG CCGTGGGAGA ACTGCTAATT GCAGAGCCGA
 89551 AGTTGGCAAG GTTCCCAAAA AACTTGATCG ATTGGGATAT GTTTTCGACA
 89601 GTGAGTGAGT CTGTAGGTTG GAGAGCAGAG GCTCCTGTAG TTAGTGAGT
 89651 GATTGCGGGA GTTGCCGACG TAGTTACAGA TGCTGGAGAA TAGGCAACAC
 89701 GGTTCGTAGT GAAGATCAAA TCTTTGATAG ATTGGAAAAC AATATCTTTT
 89751 CCGTAGATGA GGCCGCCAAG TCCTGTGAT TGTTCCCTG TAAAGTTTAT
 89801 AGAAGAAAAA TTCTTGAAGG TCAGAGTCTC TGCGGCAGAA AGTAGCGCAT
 89851 AGTTACTATT TGTAGGAGCT TGACAAGAGG TAAACGTAA GGAAAGGCCC
 89901 TTCCCTAATA AGGAGAGATT GCTGGAATA TTGATAAAGG AACTTGAAAA
 89951 ACTCGCCCCC AAGAAGTTTG AACAAACAAA ATCTGAAGTG AGTAAGATCT
 90001 CTTACCCCTG ATTCGTAATT GGAGGAACAA AGTTCCTCC GAGTCTAGTC
 90051 TCAGCAAACG CGCAACTGCT ACTACATAAA CAGGCAAGTA GACAAAAAGA
 90101 TGAAGATTTG AAAGAAAGAG GCATGCCTCA ACCCTGTCGT TAAATAAGGT
 90151 TTAAATCGT AATTGCTTC CTATATCTAG AGTATATTGA CGGGAGGTTT
 90201 TGCGAATATC ACATCCACAG TGCCCGAAGA TCTCTAGAAC ATCATTTACT
 90251 GAAGTATGGC TGGATGCTTG TACTAGCAAA GTACTTCTGG TTAGATTATT
 90301 CCCTATAGAG GTCCACGTAG CTCCATTAAT AGGTAATGTC GTATCGCAAT
 90351 CAGGATTGTG ACGATAGACA TCAGGAGCAT AGGCTACGAT TATAGTGTA

90401 AAATCTGGTC GATTATGAGA ATTTTACCA AAGCGGACGC CTACGGGAAC
 90451 TGCAACGTTG AGTAGATGAC CGTGTCCAAA AATCCTCCCC TCGGGGGTAT
 90501 TTTCTTGGAT GCCCCATGA GTCGCGTAAG CAACTTCAGC TTTTACAAAG
 90551 GGAATGATCT GCTTGAGGTT TAAGATGCGA GAAGAGAGAG TGATGGGAAG
 90601 GTTCCCTTCG AGTTCTCCTA ACCAGCAATT GTTACTCCAA GAGCAGCGCC
 90651 CTTTAGGCAA TTTTGTATAT GAAGTCTTTA CTTTCTCATT GCTACGGCTA
 90701 TAGGTAATC GAGAAGTGCC TCCTGAGAAG AATCTCGATG ATCCAAACAG
 90751 AGACTTGGTG ATGTTAGAGT ATACTGTAGC GAAATAAACG TTAGAATGAC
 90801 CGTGACCTAC GAGGTAATCC TTAGATTTTG TAAACAGCTG TCCAAAACCT
 90851 AGACTTAACG CAGCATCTGG AGTGATGCGT GTGTAGGTAT TGATGAGGTA
 90901 GCCTCCACCC ATATGGCGGT AGCTGCGTGC ATCACCGGTA TGATTGCGAT
 90951 GGAAGAAATT TGTAATTCCT GTGATGCTCA GTTGCTTCCC AGGGACATCT
 91001 TCGCCATCAG CTGCTGACGC TTGACTTACA GCTCGTAAAT CTATGACGTT
 91051 TGCCCATAGG CTATTAGGAA TGAGGGGAGC AAGACGCTCC GGATGAGGAA
 91101 GATAACCGGT TTTTTCCTCA TTTCTGTGA CCGTATGGGT TGTCGTGTCT
 91151 ATGGTAACT CCCAAGTTCC TTGATACCCA TAGGGAGATT GCTGATAGCC
 91201 GTTTGTGCCG AGACTGAACT CCGTAGTGGT TACAGTATTT GAAGTCGCTT
 91251 TGAGTTCTAA AATCGGAACT TGCTGTAAAT CTTTATTAAA CATCCCGTGG
 91301 TTGTCACAGC AATCTTGAGA GTTTTTCACA AGTCCTAAAG TTCCGGATAT
 91351 AGTGAGAGCT CCATTGGTAC TCTGCACATT AACGACAGCC GCTTTAGTGC
 91401 CATCCAAAGA ATCCAGATTG ATTACAAGCT TGTTTAAGGT GATAGCACCG
 91451 TCAGTATTAT TAGCTCCATT TGTAGTTGCT AATGTGGTCC CTGCATCCAT
 91501 GATGACGACG GACTTTTCAT CTTGCGTGAA GTTATGAACA TTAAAGGTAG
 91551 CACCGTTTCT TAAAGCGAGA GTACCGCCTT CAAGTTCTAG CTTTTGGTTT
 91601 AATGTAGATG TAGCATTTCG AGGGGTGCT GCTTCGGTAG CAGTGAGGGT
 91651 TTCTCCTGAA AAGACAATAG TCCCTGAATA CGCACCATCT GCACTGGCTT
 91701 TGGGATTGAC GACCACAGTA GCGGCTGCGG ATGCAGCAGA TAAATCATCA

91751 GATGTAATCG GATCATAGAA GTATAGGGTA TAGCCTTGCG TAGCTCCTAG
 91801 AGTGGCAAAC TTGGCATCTT TTCCGAAGTG AATACTATTG CGAGTAGGTG
 91851 TCCCCTAGT GATGCTGAGG TTCTTGTTAA AGAGGATATC ACCTTGATTT
 91901 GCGGATAGAG AGAGCTCCCC AGATTCGGGA ATAGCAATAG CGCCTCCTTT
 91951 TCCAGCTGTA TTTCCAAGAA ATATTGTAGA TTTATTAGAA TCTATAGAGA
 92001 TCTTTTTGCC ATAGAGGGCT CCTCCTTGTT CGGAGGCAAT GTTTTCTAGG
 92051 AATGTAACGC TGTTTTCTCC AGATATAGTC AGGCTAACAC CTGTTGGTGG
 92101 GGGGGTAGCT GGAGGAGTAC AGAAGATGGC GCCTCCATAT CCTAACAAAG
 92151 GAGTGAAGTC TGGTGGTGTG GGTGGAGGTG TAGGTGCAGG TAAGGAGTTT
 92201 TGTGGAGATG CTGTATTGTT TTGGAAAGTC AGGTCGCTGT TATTAGAAAA
 92251 TGTGACATTT CCGTTAGCAT AGATAGCGCC TCCTGAGCGC GAGCTATTAT
 92301 TAACGAACAA GACTCCTGAG AGGTTCCCAG AGGTGAGCAT AGATCCTCCG
 92351 GTAAGGTAAA TAGCCCCACC ATAGATCCCT GTAGCATTCG TTGAGAAAAT
 92401 CACAGGAGCG CTATTGTTGA TGAGGTTGAT CGCTGCAGAT CCCGTGAGGG
 92451 CCCCTCCATT AGAGATGGAT CCATTACCAT TAAAGAGAAG GCTCTTTTTC
 92501 GTATTTTCTA TTGTGATGCT TGTGCCTCGA ATGGCAGCTC CAAATCCTGC
 92551 AGAACGGTTG TATTGGAATA GTATGGAGTC ATTGTTTGTA AAGAGCATGG
 92601 GCGTTGTAGC GTAAATCGCC GATGCGTGAG GTATGACATT ACTCGCTGAG
 92651 GTATCTGAAG TCAAAGATTC ACAGTTATCG AAGATCATCT GACTAAATCC
 92701 TGAAAACTC AAGGGACATA GTTCAGGATT TTGGGTGATT ACACTACTAA
 92751 TCGCGGCTCC GTCAGCTGAA GAACGGATAT TTAAGAAGGA GAAAACCCCA
 92801 CCTTTTCTA AGATTTGTAG TGCTCCCGCC CTATTGCTAA AGCAACTGGA
 92851 AGAGGTTCTG GATATGGCAT TATCAAGATT CGCAATGTAG AGATCCCCTG
 92901 AAAAAATACA GAGTGTCCCT CTAGGATCAG AAAGTGTGTG GTAAGGAAAA
 92951 ATCTTCCCAC TCGATCCATC AAAGTTCTCG GAAGGCATGA TAACTTCTAC
 93001 AGTAAACGCT GTTGAAGCAA AACATGGCGC CAGTGTGGTA GAAATTAAGA
 93051 ACTTACGAAT AGACGTTTTC ATTTGCACGT AGAGATGAAA CCAGATTATC

93101 CTACAAATAA GGGAAAGGCT GTAAAAAAC AAGTACAATA AGACACAGTT
 93151 TTAATCTCTT AATTTTGACA GCTTTAAGAT TACAGGATAT TTAAAGGGC
 93201 ATTTTCCCAT TTCTTACATT GCTTTCTTAG AAGAATACTT GATAGAAAAT
 93251 GGCGATTCTA TTTTGTAAAA ATCTCAAGAA ATTCTCCCAA ACGAAGATGT
 93301 TTTAGAGAAC CTGTAAAGTA GAAAATGGCG CTACGCCCCG GACTCGTGAA
 93351 CATGTGTGCA TAGAGGGACC TATCTCTGAA TACAGATAGG TCCCAAATC
 93401 TTCTTAAAGG GGATTCCCTT AATTATAGTG TAGACTTAGA GATTATGGCG
 93451 TAGAAGTGAT CTCAGGATTC CAAGTGAGGA AAACAGTTTT CTTAGTTGGT
 93501 GTGATCCTTC TGTCTTTATC CAGAGAAGGG GGGTACTCTT CCCAGCTTAG
 93551 GGACCAGAAA CCTCGAACGG CGTCATTCTC TGTGTGACT TGGGTTCTCT
 93601 CCAAAGTGGT CTCCCCTAGG AGAAGACTGT CAAAAGAAGG CCCTAGAAGT
 93651 TCAAGAAGAG GAATCGTAAA GGCTTCCTTA AATTGAGGAA AGTCATAGAC
 93701 GGACCAATAA GCCTCAAAGG CAACTTTAAG TTTTCTAAT GAGACAAGAG
 93751 CATCCTTGTC CTCGGCGCGA ATCCTTACAG GAACAAAGTT GTCGGTATCT
 93801 TCAATCAGGA TGTGCAGATT CTGAACCCGA GCATCTCCTG AGCAGAGCAG
 93851 AGTGGTTCCT GGAGACATAG TAAGCGTAGA GCTTGCTTCC TGCTTAAAAG
 93901 AATGCAGTTG CAAGGTAACC CCATCCGATA GAGAGAGAGT TCCTCCTGCT
 93951 AATGTGACAT CTTGTAGGAT TGTGGAAGTA AGATTTTCCG CACAAACTTC
 94001 ATGATCATCC AGGCATAGTC CTGAGAAGCT AATTGTTCTT TCATAAGTTT
 94051 CCTTTCCTTC AGGAGCATTG ATTACAAGAT CTGTAATTTT ATGCGACTCG
 94101 CTATGGCTTA TAGGATCATA GAAATAAACT CCGGATTCTG AAACAGCACG
 94151 TAGGTTCTTA AACTGTGCTC CAGATTGCAG ATGGATGGAG TTGTGTATTG
 94201 TATTTCCGTC TTGTGATGCT GTATTTCTT TGAAGATGAG ATCTCCGCTT
 94251 TTCACGGATA TAGAGATCGA TCCTCCAGGA GCAATGGCAA TGGCTCCTCC
 94301 ATTACTATTC ACGTCATGAT AAGCATGATT ATTTTCAAAA CACGAAGGTC
 94351 CTCGAGTCGT GAGTGTGAGA TTGTGGGTAG ATATGGCGCC GCCATAACCT
 94401 TGGCTCACAT TGTCTCTAAA CACCAGGTAG CGGTTCCCGC TGAGATTTAC

94451 TGAAGGACGA CTCGCCTTAG AACCTAAAAG GTAGGGAGTA TAAATCGCAG
 94501 CTCCACTCCA CGAAGAGTAG TTCCCACAGA AAGTCACTTC CTCACTATTT
 94551 TCGATCATCA CGGAACCAAG ACTATAAATC GCTCCTTGTC CTTGAGGTAG
 94601 TAGAGGTGCT GAGGTGAACG CTAAGTAAGA AAAATTAGAG AGAGTGAGAG
 94651 TGGTGTCTCC AACGCGGTTT GAAATGGCAG CGCCAAAACC CTCGGTCATA
 94701 AGGTTGTGAA AAGTGAAGTT GCAACGGTTG CCCATGAAAA AAAGATTTCCC
 94751 AGATCGATTT ATAAAAACCC CAGCATCTTC TTGATCATGC TTAACGTTGG
 94801 AAATCCTCAC GTCATCTAGA AAGATGTAAG AAGTTCCTTC TGGATAACAG
 94851 GTAATTTTAG GTTCTAAGCT TTTATTATTG ATAGCACCGT TATAACCATC
 94901 ACTTTCATGA AGATATACAA CTTGTGCTGC TGCAGGGAGA GCGAGGAATA
 94951 AAGCCGAGCA GGTAAGAAAA TTTCGAAGTA TGGTCATGGT TTCCTCGTTA
 95001 AATCAATAAG GTTGAAGCAA CTTTAATAAA CAAGAAAAAA AGAAGTCAAT
 95051 AAGAATAGAT TATTGTCTAT TAATTATTTA ACTGTTTTTA AAATAAAATT
 95101 ATAACTAGAA ATTATTAAAA GAAATCTTTT TTGAAGAGGG ACAAATGTTA
 95151 TTTTTTACAG TTTGCAAGGA AAGCATTCCC TATAGCAAAT ATTTCCCTAA
 95201 AAGTATGAGA AACTCCCTA GAAGAACTAG GGAGTTTTAG CAATCTAGAA
 95251 TCGGAGTTTG GTACCAACAT CTACATTGTA GTTCCTTGAA GATCCACGGA
 95301 GTTCCATAGC GTAATGTCCG AAGAGCTCAC AATTGGAGTT GTAGACGTAG
 95351 TTGTTGCTAC CCCTCAGTAA AAATGCCTGT CTTGAAAGAT TGCCACCGCG
 95401 AATTTTCCAA GAGTCTGGGC TCATCACAAG AGTCGCTGTA GATTGGGGAT
 95451 TGTTACGATA GACATCGGAA ACAAAGAATC CTGAGAGATC ATAGGTGTAG
 95501 GAATCTCCGA TATCCCCTG CACGAATTTT GCACCCACAG GAATCGAGAG
 95551 GTTAAGCAGC CTTCCAATAC TAAAACCACG GCCATCACTA GAGCTTTTGA
 95601 AGAAGCTATT TTGTGATACA TAAACCATTT CGACTTTCAT CTGTGGAATG
 95651 AAGGTCTTGA AAAGAGGATG TGGGTTGGAA AGAACAAAAG GAAGGTCTAG
 95701 GCCGATACCA CCAGCTATAC ACTCGTTGCT CCAAGAACCT TCGGATTCTG
 95751 GCAATGAGGT ATAGTGCCTT TCCATACGGT TGTCTGAATG GCTGAACGAA

95801 ACTTGGACAT CCAAGGCTAG GGGAAATTTCC CTAGGGAATT TTTCTATAGC
 95851 TGATTTCAGAA AACTTTGCTC TTCCTAATCT CAAATAGTTT TGGGGTTGTA
 95901 GGGTATGAGA GTGCTTGAAG AATAAAGTTC CACCGTAGGT TCTAGAGTTG
 95951 TTGTGAGCGA TAAAACAATC TTTGTCTCTA GCAAAGAGAT GGCAGAACGC
 96001 AAAGGTAAAT AGGTCGTCTT TAGGAGTGTG AGCACTTCCA CCGATGACGT
 96051 AGCCTCCAGA GGTATGACGG AAGCCTTTGC GATTTTCATC TCCAGTCTTA
 96101 TGCAGGAAGT TCGTCATGGA GGAAACCCAG AAACCTTGTT TGTGTTCCAT
 96151 ACCAGTTGCG CCGATCTCTA CAAGCTGTTG CAGAGAGCGA ATGTCAGTAA
 96201 AGACTCCCCA TAGGGTATTG CATACTAACG CAGATTTTCT TTCGGGGCTG
 96251 GGAACAAATC CTGTTTTGGT CCAAGTTGCC GTGGCCTCTT TTGTATTTGT
 96301 AGCTGTATCC GTAGTCCAAT TAACATTCCA TTGTCCTTGG AATCCGTATT
 96351 CTGAATTAGG ATCCTCAGCA GGAACAGGGA TAAGGCTGCT GATGTCAACG
 96401 TTAGTATCAA CATCAGCATC AACCGTGATT TTTAATAGAG AGAAGAGCTG
 96451 GTCATGGCTG AACATATGAC TTTCATAAAT GTTCCCTTCA ATATCAATCA
 96501 GGTGAGCTT CCCAGATACG ATCACTTTAT TTGAAGCACC TTTTGCTGTT
 96551 AGGCTGACGG GCTGCTTAAG ACCTAAGGAG TCAACATTGA TTCCTAGGTT
 96601 CGTGATTGTA ATACTCCCAG CTGTAGTTGA TAATGTCGTT CCTGAATCCA
 96651 TGCCGAGGAG AGAACCGGCC TCTTGAGAGA AGCTCGTGCT CTCTAAAGTG
 96701 ACTCCCTTTT GTAGCAATAA CTTTCCTCCG GATAGGGAGA CTGGCTGCGT
 96751 GAATGAAGAT TTTAAATTGT CAGCAACTTT AAGTTCATCT GCTGTTAGGG
 96801 TTTCTCCAGA AAATAGAATC GTTCCTTGAT ATGGATTGAG AGCTCCCGCA
 96851 GAGCCGTTAT TTATCTTCAA TACGTCTGAT GAGGTTCCCTT CTGAAGTGAT
 96901 GGGATCATAG AAGAAAATTG TATGATTTTT AGCAGCCCGT AATCCGTGA
 96951 ATTTCCCGTT ACTTCCTATG TTGATCGCAT TACGTTTAGG AGTATCGGTA
 97001 CTTCCGGTTG TTGTAAGGGT ATTTCTTACA AAGGTAATGT TTCCTGTCTC
 97051 TGCAGAAAGA CTGAGCTCTC CTGAGGCATC GATGCTGATA GCACCCCCCT
 97101 TAGGAGTTGC TGATGAGACA TTATTTTCGTA GAAACTCTGT AAAGCCTCCA

97151 GAGGAAAGGG CTAGCTTTTT AGCATGGATG GCGCCACCGC TTGTTTCTGC
 97201 TACGTTTGAA GCAAAGATCA GAGTCTTATT GTTAGAGATT ATCAGTTCAG
 97251 GAGATCCACT CGCCTTGGTG TTGCAGATCG CACCGCCAGT AGTTTTTCGCT
 97301 GCATTCCCTT CAAAATATAG AAATTTGTTG TTCGATAGTA TCGACGTGCC
 97351 TTCATCATCG ATAGCGCCTC CTGACGTAGA CGCTATGTTA GATAGGAATC
 97401 TAACATAACC TGTGTTATTT GCTATGCGAG CGCCTGCTGT AGTAGCAATT
 97451 GCTCCTCCCT TTGTTGATGA AGAGTTGTTA CTAAAAAGAG CATCTCCAGA
 97501 AGTGCCAGTT AAAAGGAAAG ACGCTCCTTT GATAGCTCCA CCATCTGCAG
 97551 TAGAAAAATT CCCAGCAACT ACAAGTTTAC GAATATTTTC TAAATTTACG
 97601 CCTCCTGCTG AGGAAAGCGT TCCCTGACCT GTAGTAACCG TTGTGCTAGG
 97651 AGAGGAATCA AACTCAGTA AGGAAAACCC TGAGAAGGTA AGATTCTTAT
 97701 TTGCTGTTGT AGATGCAGCA GCACCTGCAT GAGTGCCAGC ATCTATAAAG
 97751 CCAAACGTTA AGCTATGACC GTTCCCCAAG AAGGTAAGAT TGTCCGTGGT
 97801 TTGCTTAAAA CAACTGTCAG ATAAGGGAGT GCCTTTTCCA GGCTCGTAAA
 97851 AGAAGACATC TCCTGTTAGA GAATATGTTG TGGCTGAAGT TTTTGGAGTA
 97901 AACGTTCCCTG AATCGATATT TCCATTAAAG CTATCATCAG GTGATAAAAG
 97951 TTCCTCGTTA GCTAGTGA CTAGGTGACA TGAGAAAGCT AACACGGAGG
 98001 AACTTAAAC CCAAGGAATC GAAGTCTTCA TGGAATGCT TTTGTTTTTT
 98051 AGAGAACTAT TCGCATCAAT ATAGAAACAA AATAAGTAAA TCAAGTAAAA
 98101 GATGACAAAA CAGCTGTCAA GAATTTTTAT CTTGACTCTC TGAGTTTTCT
 98151 ATTTTATATG ACGCAAGTAA GAATTTAATA ATAAAGTGGG TTTATGAAAT
 98201 CGCAATTTTC CTGGTTAGTG CTCTCTTCGA CATTGGCATG TTTTACTAGT
 98251 TGTTCCACTG TTTTGTGCTG AACTGCTGAA AATATAGGCC CCTCTGATAG
 98301 CTTTGACGGA AGTACTAACA CAGGCACCTA TACTCCTAAA AATACGACTA
 98351 CTGGAATAGA CTATACTCTG ACAGGAGATA TAACTCTGCA AAACCTTGGG
 98401 GATTCGGCAG CTTTAACGAA GGGTTGTTTT TCTGACACTA CGGAATCTTT
 98451 AAGCTTTGCC GGTAAGGGGT ACTCACTTTC TTTTTTAAAT ATTAAGTCTA

98501 GTGCTGAAGG CGCAGCACTT TCTGTTACAA CTGATAAAAA TCTGTCGCTA
 98551 ACAGGATTTT CGAGTCTTAC TTTCTTAGCG GCCCCATCAT CGGTAATCAC
 98601 AACCCCTCA GGAAAAGGTG CAGTTAAATG TGGAGGGGAT CTTACATTTG
 98651 ATAACAATGG AACTATTTTA TTTAAACAAG ATTACTGTGA GGAAAATGGC
 98701 GGAGCCATTT CTACCAAGAA TCTTTCTTTG AAAACAGCA CGGGATCGAT
 98751 TTCTTTTGAA GGAATAAAT CGAGCGCAAC AGGGAAAAA GGTGGGGCTA
 98801 TTTGTGCTAC TGGTACTGTA GATATTACAA ATAATACGGC TCCTACCCTC
 98851 TTCTCGAACA ATATTGCTGA AGCTGCAGGT GGAGCTATAA ATAGCACAGG
 98901 AAAGTGTACA ATTACAGGGA ATACGTCTCT TGTATTTTCT GAAAATAGTG
 98951 TGACAGCGAC CGCAGGAAAT GGAGGAGCTC TTTCTGGAGA TGCCGATGTT
 99001 ACCATATCTG GGAATCAGAG TGTAACCTTC TCAGGAAACC AAGCTGTAGC
 99051 TAATGGCGGA GCCATTTATG CTAAGAAGCT TACTACTGGCT TCCGGGGGGG
 99101 GGGGGGTATC TCCTTTTCTA ACAATATAGT CCAAGGTACC ACTGCAGGTA
 99151 ATGGTGAGC CATTTCTATA CTGGCAGCTG GAGAGTGTAG TCTTTCAGCA
 99201 GAAGCAGGGG ACATTACCTT CAATGGGAAT GCCATTGTTG CAACTACACC
 99251 ACAAACTACA AAAAGAAATT CTATTGACAT AGGATCTACT GCAAAGATCA
 99301 CGAATTTACG TGCAATATCT GGGCATAGCA TCTTTTCTA CGATCCGATT
 99351 ACTGCTAATA CGGCTGCGGA TTCTACAGAT ACTTTAAATC TCAATAAGGC
 99401 TGATGCAGGT AATAGTACAG ATTATAGTGG GTCGATTGTT TTTTCTGGTG
 99451 AAAAGCTCTC TGAAGATGAA GCAAAAGTTG CAGACAACCT CACTTCTACG
 99501 CTGAAGCAGC CTGTAACCTT AACTGCAGGA AATTTAGTAC TTAAACGTGG
 99551 TGTCACTCTC GATACGAAAG GCTTTACTCA GACCGCGGGT TCCTCTGTTA
 99601 TTATGGATGC GGGCACAACG TTAAAAGCAA GTACAGAGGA GGTCACCTTA
 99651 ACAGGTCTTT CCATTCCTGT AGACTCTTTA GGCGAGGGTA AGAAAGTTGT
 99701 AATTGCTGCT TCTGCAGCAA GTAAAAATGT AGCCCTTAGT GGTCCGATTC
 99751 TTCTTTTGGA TAACCAAGGG AATGCTTATG AAAATCACGA CTTAGGAAAA
 99801 ACTCAAGACT TTTCATTTGT GCAGCTCTCT GCTCTGGGTA CTGCAACAAC

99851 TACAGATGTT CCAGCGGTTC CTACAGTAGC AACTCCTACG CACTATGGGT
 99901 ATCAAGGTAC TTGGGGAATG ACTTGGGTG ATGATACCGC AAGCACTCCA
 99951 AAGACTAAGA CAGCGACATT AGCTTGGACC AATACAGGCT ACCTTCCGAA
 100001 TCCTGAGCGT CAAGGACCTT TAGTTCCTAA TAGCCTTTGG GGATCTTTTT
 100051 CAGACATCCA AGCGATTCAA GGTGTCATAG AGAGAAGTGC TTTGACTCTT
 100101 TG TTCAGATC GAGGCTTCTG GGCTGCGGGA GTCGCCAATT TCTTAGATAA
 100151 AGATAAGAAA GGGGAAAAAC GCAAATACCG TCATAAATCT GGTGGATATG
 100201 CTATCGGAGG TGCAGCGCAA ACTTGTTCTG AAAACTTAAT TAGCTTTGCC
 100251 TTTTGCCAAC TCTTTGGTAG CGATAAAGAT TTCTTAGTCG CTAAAAATCA
 100301 TACTGATACC TATGCAGGAG CCTTCTATAT CCAACACATT ACAGAATGTA
 100351 GTGGGTTCAT AGGTTGTCTC TTAGATAAAC TTCCTGGCTC TTGGAGTCAT
 100401 AAACCCCTCG TTTTAGAAGG GCAGCTCGCT TATAGCCACG TCAGTAATGA
 100451 TCTGAAGACA AAGTATACTG CGTATCCTGA GGTGAAAGGT TCTTGGGGGA
 100501 ATAATGCTTT TAACATGATG TTGGGAGCTT CTTCTCATT CTTATCCTGAA
 100551 TACCTGCATT GTTTTGATAC CTATGCTCCA TACATCAAAC TGAATCTGAC
 100601 CTATATACGT CAGGACAGCT TCTCGGAGAA AGGTACAGAA GGAAGATCTT
 100651 TTGATGACAG CAACCTCTTC AATTTATCTT TGCCTATAGG GGTGAAGTTT
 100701 GAGAAGTTCT CTGATTGTAA TGACTTTTCT TATGATCTGA CTTTATCCTA
 100751 TGTTCTGAT CTTATCCGCA ATGATCCCAA ATGCACTACA GCACTTGTA
 100801 TCAGCGGAGC CTCTTGGGAA ACTTATGCCA ATAACCTAGC ACGACAGGCC
 100851 TTGCAAGTGC GTGCAGGCAG TCACTACGCC TTCTCTCCTA TGTTTGAAGT
 100901 GCTCGGCCAG TTTGTCTTTG AAGTTCGTGG ATCCTCACGG ATTTATAATG
 100951 TAGATCTTGG GGGTAAGTTC CAATTCTAGG AGCGTCTCTC ATGTCTCAGA
 101001 AATTCTGAGA GAGATCGCAT TTAGGATTTT CTAAACACG ACTCACCTTG
 101051 TTTTGAACC AGGAGAGATC GGGGATTAAA AAGGCAAGAG GGCAGAGTTC
 101101 GTGAGGTCAC G TACTCTGCC TTTCTTGTTA CAAACACGTT TAAAAATTAA
 101151 GGAAATTTTT TAATAGAAAC CCGTTCTTTA AAATACGTTT CTTTAATTCT

101201 TATTGAATAA GATAATTCAC TATTTTTAGA TCCTAAATTT TAAGTGGTTT
 101251 TTGTTATGCT TCTTATAGAG AATAGCTGCA AAGATTAGAG TTGCAGAGAC
 101301 GGTACGTCTC TTTCTTTTTT AAGGGAAGGG GTGTTGTTAC ACCCATCCTA
 101351 AGATTTGTGA GATTCCCCTC AGGCAGTAAC TTTTACAATC GTACTTTATG
 101401 TTTTGATCTA GCTGTTTTCT TGTCTTTAAT TTATTCAACC ATCGAGAAGA
 101451 GAGATCCATG AGTGGAAATG TATTTTATTA GGATCATCTC TAAGGATGGA
 101501 AATGATGAGC CCATTCCAAC AACCTGAGCA ATGTCATTTT GATGTTGTGG
 101551 GAAGTTTCTT ACGTCCTGAA AGTCTTACAC GAGCACGCTC TGATTTTGAA
 101601 GAAGGAAGAA TTGTCTATGA GCAGATGCCA GTTGTCGAAG ATGCTGCTAT
 101651 TCGTAATCTC ATAAAAAGC AACAGAAGC AGGTCTTATC TTTTTTACTG
 101701 ATGGGGAATT CCGTAGGTAT AGTTGGGATT TCGACTTTAT GTGGGGATTC
 101751 CATGGCGTGG ATCGTCGCAG GGACTCTAAT GACCCTGAAA TTGGAGTGTA
 101801 TCTTAAAGAT AAAATCTCCG TATCAAAACA TCCGTTTATA GAACATTTCCG
 101851 AGTTTGTCAA AACTTTTGAG AAGGGAAATG CAAAAGCAAA ACAAACGATT
 101901 CCTTCTCCAT CACAATTTTT CCATGAGATG ATTTTGTGCTC CTAATCTGAA
 101951 AAATACTCGG AAGTTTTATC CTACGAATCA AGAGCTAATT GATGATATTG
 102001 TCTTTTATTA TCGCCAAGTC ATCCAAGATC TTTATGCTGC AGGTTGTCGT
 102051 AATTTGCAGT TGGACGATTG TGCTTGGTGT CGCCTCTTGG ATATACGAGC
 102101 GCCTTCTTGG TATGGTGTG ATTCTCATGA CAGGTTGCAG GAAATTTTAG
 102151 AACAGTTTTT ATGGATCCAT AATTTAGTGA TGAAGGATAG ACCCGAGGAT
 102201 CTTTTTGTA GTCTGCATGT CTGTCGTGGT GATTATCAGG CCGAGTTTTT
 102251 CTCTAGACGA GCTTATGATT CTATAGAGGA GCCTTTATTT GCTAAGACCG
 102301 ATGTGGATAG TTATCACTAT TATTGGGCTC TTGATGATAA GTATTCAGGA
 102351 GGTGCTGAGC CTTTAGCTTA CGTCTCTGGA GAGAAACACG TCTGCTTGGG
 102401 ATTGATCTCC AGCAACCATT CTTGTATTGA AGATCGAGAT GCTGTGGTTT
 102451 CTCGTATTTA TGAAGCTGCG AGCTACATTC CCTTAGAGAG ACTTTCTTTG
 102501 AGCCCGCAAT GTGGGTTTGC TTCTTGTGAG GGAGACCATA GAATGACTGA

102551 AGAAGAACAG TGGAAGAAGA TCGCCTTTGT GAAAGAGATT GCTAAAGAGA
102601 TCTGGGGATA AAGAATCCGG AGTTTTTATC GACTCTAAGA GTTTTCGGAT
102651 CATAGAAAAC ATTTAAATAT TCAAGAGTCT TTGGCTATTG GATCATAGAC
102701 AGTCTTAGTA TACTAAAAAG TCTTTGGATT CTAAGACGGG CAGAGTTCGT
102751 GAGATCACGT ACTCTGCCCC TTCTTTCTTG TGATCTAGCG ACTTCTTTGA
102801 ATCTTCGACC TCTTGTAATC TGGGATTTTT TCTAGTTCTT AGATTCCTCT
102851 GATCTTTCGA CTTCTCCTCG TCTAAACAAG GCGCATTGTC TTTGAGAAGT
102901 CCCTAGATAC ACTCAGGATC TCTTAGAATT TCTAAGGGAT CAGGAACGCT
102951 TTTAGAACTG GAACTTACCT CCAAGATCTG CATTGTAGCT GCGTGAAGAT
103001 CCACGAATTT CCATAGATAG GTTACTTGTG ACCTCAAGAT TTGGAGAGAA
103051 GGCATAAAAG ATCCCTGCTC TTCCGATACC AGCTTGTCTT GAGAGATTCG
103101 TTCCTGTAGT TTTCCACGAG GTATTGTTGA TTAGGAGAGC TGTCGTGCAG
103151 TCAGGATTCT TACGATAGAC ATCGGCAACG TAGATGACAG TAGCTTCGTA
103201 AGACGCACGC TCGTTTCTCG AGAATCTCTC GAAGGTAATT CCAATAGGCA
103251 CAGAGACGTT AATTAAATCA CCGCTATCGA AAGATCGTAC CAAGGTAGTA
103301 TTACGTTCTT TGAAGCTATC TTGGTGTATG TACGAAGCTT CTACTTTGAT
103351 GAAAGGAAAA TACGCGTGGA AGAGACCCTC ATGGCTTAAA GCAGTGTGTG
103401 GTAGGGAGCT CGCAAGTTCC AGAGCGCAAC CGTCATTATA CCACGAGCTC
103451 TCTCCCTTTG GTGCTTGGGT GTAATAGGTT TTCATAGTAT TTTTACTATA
103501 GATATAGCTG ATCTGAGCAT CAAAGAGGAC AGGCTGCTCA CTTTCAGATC
103551 CAGGAAGGTA GCGTAACAAG CTTGGAGAAG ACAAGGTCGC TAGATGCTGG
103601 AGATGGAGAG AAGCTGCATA GGCAGAAGCT CTATTTTAT TTATAAAGTG
103651 ATCTCTATCT TTCCCGAATA ATTGGCAGAA GGCTGCACTG ATAAGATTAT
103701 CAGAAGCTAA TGTGTAGTC GCTCCTACAA CATAACCTGC ACTTATGTGG
103751 CGAAAACCTT TATTTATCTT CGTGCTATCT TTATGGAAGA AGTTCGAGAT
103801 CCCTTCACAC CAGATGCCGC GAGTTTCTTG AGATTGGCGT ACTTTAGTGG
103851 CTACAAGCTG TTGTATGGAG CGCACATCAA CAAAGGATCC CCATAGCGTG

103901 TTAGCAACTA AGGTTCCACG ACGCTCAGGA TTCGGATTGT ATCCTGTTTT
 103951 TGTCCAGGTA AGAGTCGCTG CTTTGGATTT AGTCGCAGTA TCCTCTTGCC
 104001 AAGATAATGC CCAATTCCCT TGGTATCCCC AATGGATAGG ATTTTTTTCT
 104051 AGGGGATCAG CAGCTAAGTC TGTGATGTGA ATATTGCGG GGTGTCAGC
 104101 AGTAAGAGTG AGACAAGAAA AGACTTGAGG GTTATTCCAA GAGACATCTT
 104151 CGTAGACATT TCCAGAAGGA TCTACAAGAG AGAGCGATCC AGATAAAGTG
 104201 ACTGTCTGAC TTGCTTGTGT TGCTTTTAGC GTAGCCTTCT TGGTCTCTTT
 104251 TAAGGAATCT ACATTGAGAA CAAGATTATT GATAGTGATC CCATCAGCGG
 104301 TTTCTAATGT GGTCCCTGCA TCCATGAGGA GGGTAGAGCC CGGAGATTGC
 104351 GAAAAGGACT TAGCAACTAG AGTGACTCCT GATTTAAGAG AGAGTTGCCC
 104401 TCCCGCAAGA GTTAGAGGTT GCTGAATTGT AGATTGAGA TTATCAGCTT
 104451 CTGCAGCTTC TGCTTCCGAG AGCTTCTCTC CAGAAAATAC GATGGTTCCT
 104501 TGATATGCAG GATTCCCTGC AAGGTCAGGA CCATTTAAGT TTAGAGCATC
 104551 TGAGAGAGCT GCAGTGATGC TAGTTGTTAT AGGATCATAG AAGTAGATAG
 104601 TATTGCCTTG AGAGGCTCGC AGCTGTACAA TCTTAGCATT GGTGTTTCCG
 104651 ATGTTAATAG AATTTCTGGT AGTGGTCTGA CTCGAAGAAG CTCCTTTGAC
 104701 TACTGTGTTT CCTTCAAAG TGATGTCTCC ACCAAGAGCC GAAAGACTCA
 104751 AAGATCCAGA GTCAGCAATC GCAATTGCTC CTCCTAAGGG AGCTGCAGTA
 104801 TCTATAGCAG AGTTGTTTTT AAAAAGCGTA GGTCTCCAG AAGAAAGAAC
 104851 TAGATTGTCA GTATAAATCG CCCCACCACT AGTAATTGCT GTATTCCTA
 104901 TAAAGTTCAG TTCCCCGTG TCTGATAGAG TTAAGACTGG TTTGGGGGCT
 104951 GATGTACTAC TACAGTAAAT GGCTCCCCCT GTAGCTGAGG TTGCGGTCAC
 105001 ACTATTGTTT ATAAAGCTAA TTGCTTTGTT GCTGCTAATA AACTGCTAG
 105051 CTTCCGTGTA AATGGCTCCG CCATTGTTCG CCGCGGTATT TTCAGAAAAT
 105101 GATGCTGAGT TTAACGTATT GTTAATTGTA ATCCCTCCCG TGAATAGAG
 105151 GGCACCCCT TTTTGCGTTG CTTTGTTTTT GGCAAACGTT AGGTGGGGT
 105201 TTAGCGATAG ACTGATAGAG CTGCCTTGA GGGCGCCTCC ATTGTCATTA

105251 GAAAAGTTTT GGCCAAAGTA GCAACTATAG TTCGACTGAA TAGAACAAGC
 105301 TCCTGTGGAC TTGATGGCTC CTGTTCTGT GGTAGCATTG GTGGTTTGTA
 105351 TTAGTGACAA ATAGGAGAAT CCTGAAAAGG AGAGAAGCTT ATTTGCAGCT
 105401 GTATTGGTAA AGGTACAGTT CGCTCCCGCA TCGATATTTT GTAGGAGAAA
 105451 TTGGTAGCCG TGGCCTTGA AAGAAAGATT CCCAGTAGTT TCTTTAAAGC
 105501 AGGAAGCGGT TAGAGCTGTC GGAGATCCTG CATTGGTGAT TGAGACATCC
 105551 CCTGTTAGAT TATAGATAGT TCCATCTGCA TTTGTTGTTT GGGCTGGAGG
 105601 AGTGTAGGTT CCTGGTCCAG AGAAGCTATT GGTAGGTCCT AGATTGATTT
 105651 CAACAACAGC AGCAAACGCA GAGAAATTTA GTGACAAGGG AAGTGCTAAA
 105701 GATGACGAGA TTAAAAACCA ATGAAGAGAG GATTTCATGT AGAGGGCTAT
 105751 AGGTGGTTTA ACAAATTATT TCACCACATA CTGCAATAAA TTAAAGAAAAG
 105801 CAAGAGGAAA GGAGAGACTA GTAAGTTAAG AATCTACAGG GTTTTTATAA
 105851 GAATTCCTCC CTAAAAGTTT AGGGAGGAAA GTAGGAACTA GAATGAGTAT
 105901 CTTAGCCAC AATCTACATT GTAGATGTGT GCTGAGCCAC GAAGCTCATA
 105951 AGCAGCTTCC CCAGAGAGTT CTACATGAGG GGAGAGAGTC AGATGGCTTC
 106001 CAGCACTTGC TAAGAAGGCT TGTCGTGCGA GGTTTTTACA TAGCGAAGTC
 106051 CAAGAGGCTC CACTGACCAT TAGAGAAGTA CGCGAACGGG GATTTTTACG
 106101 ATACACATCA CCAATGTAGG CTAGAGAAAT CTCGAAATTA TTTTTTTCAT
 106151 CTTGGGAGAT TTTTCTAAC CGAATGCCGA CAGGGATAGA GCAGTTCCT
 106201 AGGTCTCCAT CATCAAAAGC ACGGGCTTCA GCGCCACTCT CTTTAAAGTT
 106251 TTGTTGGCGG CTGTAGACTG CCTGGAACCT TAAGAAGGGG AAATATCCCT
 106301 GGAAGAACGG TGCTTCTTTA GGGAGATATA GAGCCAGAGA TCCTCCGAGC
 106351 TCTAGAGCCC CAGAGTTATT GGTCCAAGAG CCTTGAGCTT CAGGATAGGA
 106401 AGTATAGCGA GTATCCATAT CATTTTGTAG GTAGCTGTAG CTTAGCTGGG
 106451 CATTCAAAAT GAGAGGAATA TCTTTCAGCA TGTCGGTGAT ACTTCCAAAT
 106501 GAGGGCATGG GAAGTCCTCC TAGGAATGCT CGATGTTGCA GGTATAGCGA
 106551 CGCTAAATAG TTATGAGAGG TATTTTCAAC TATAAACAGG TCTTTATCTT

106601 TACCGAAGAG CTGGCAGAAA GCTACACTGA AGATATTTTC AGAAAAATCT
 106651 TCAGCACTTC CTCCAACAAT ATAGCCGTAG CTTTATGTC GGAATGCTTG
 106701 GTTAGTTCCT GATTTATCCT TATGGAAGAA ATTTCGAGTT CCTGATGCCC
 106751 AGAGTCCTCG TTGCTGATAG ATACTATTCG CTTGAGATGT CATGATCTGC
 106801 TGTAGAGTGC GAATGTCAGT AAAGGATGCC CATAATGAAT CGGGAACCTAC
 106851 GGAAGCTCTA CGCTCAGGAT TAGGGTTGTA GCCCCGTAGTT ACCCAAGTCA
 106901 TAGTTCCTGA TTTTGCAGTT GATGTGTCTG CCCAAGTGGC TTCCCAATGT
 106951 CCCTGATACC CGTAATGAGG TTCTGGAGTT TGTACTGGAG AAGTGAGAAG
 107001 CGCATCGATA TAAATATCGC TAGCAGCAGT AGCAGCAGTG AATACCACCA
 107051 AAGGCTGCGT GAAGGCTTGG TTTATCGTAT GGCTTTCATA AAAATTGCCG
 107101 CTACTATCTT GGAAAACAAG AGGAGAGGTT AGAGTTATAG TTTTGTGGC
 107151 TCCTGCTGTT TCAATGGACA CACTCTTATT TCCCTCTAAG GCAGAAAGAT
 107201 CAACGACAAG TTTGGTAAGA CTGATAGCTT CAGTATCTGC TTTGAGCTTT
 107251 GTTCCTGGTT GCATGAGGAG TGTAGAGCCT TCAGTCTGTG TGAAACCATT
 107301 GACATCTAAC TCGACATTTT CTTTGAGTGC TAAGGTTCCA GAGGCTAGAG
 107351 CCAATGGTTG CTTTAATATA GATGTGAAGT TATCAGCAGC TTTCGCTTCA
 107401 TCTGCAGAGA GCTTTTCCCC AGAAAATACA ATCGTTCCTG AATAATCTAA
 107451 AGGCGAGTTG CTATCCGGTT GGTGATGGT CAGAACGTCT GAAGCTCCTG
 107501 TGGTGTTAGA TGCAATCGGA TCATAGAAAT AGATAGATTG GCCTTGGGCT
 107551 GCCCTTAAGT TCGTAATTTT TGCTGACGAT CCCAGGTAGA TAGCATTCGG
 107601 TGTCGATGTT GGCGCGGAGG TTGAGGTTAG AGTGTGCCA AGGAACGTGA
 107651 TGTCTCCTTG ATTTGCAGAG AGACTTAAAG ATCCAGAGTC GGCAATTGCA
 107701 ATAGCGCCGC CCTTGCCTGC AGCTGTGTTT CCGCATCTAT TATTTGAAAA
 107751 TAGGGTAGGG CCAGCAGCGG AAAGATCTAG ACCATGGGCA CAGATTGCTC
 107801 CGCCTTGAGT TACTGAAGAG TTCTCGGCGA AGGTCAGACT TTTATTTCCA
 107851 GAGATAGTAA GAGTAGGAGT CTCTCCTGTT TTTTACAAT AAATGGCCCC
 107901 GCCCTTGCCT GCAGCATCTG TTGCAGTGTT TCCAGAGAAG AAAAGGGAGC

107951 TATTTTGAGT AATCGAGGAG CTGGCTTCAA AGCCCAGAGC CCCACCCCCA
108001 GTTCTCCTT TATTATTCAT AAAGACTAAC TGGCCGGTGT TTCCTGAAAT
108051 ACTTGCAGCC GCAGAGCTAT AGATCGCTCC ACCTAATTTT TTTGCGCTAT
108101 TACTAGTGAA GGTTATAGAA GAGGTATTCC CAGAAATAGA AAGAGTTTTT
108151 GTGGTGATCG CTCCGCCATT GTTATTAGCT TCATTGGAGA CGTTTTGGCT
108201 AAAGAGAATC GTTCCATTAT CGGTAAGATT TAAGGCTCCT GCAGAACTTA
108251 AAGTACTTTT TCCTGAAGCA ACTGTAGTTC CAGGAGCTGC AATGAAGGAA
108301 AGGTTAGAAA ATCCTGTGAA TGTTAGGGCT TTATCAGCAG TTGTGCTTGC
108351 CGCAGCTCCT GCATTGGAAC CCGCATCTAC CGTGTTGAAT GAAAATGAGT
108401 ATCCCTTTCC AGTAAATGTC AGATCACCCG TAGTTTCTGT AAAGCAGCAG
108451 CCTGTTAATG CTGTGCCTTT CCCAGCATCG TTTATATAGA CATTTCCTGA
108501 TAAGACATAG TTCGTTCCAT TGGCATCTGC TGTAGATTTT GGAGTAAATG
108551 TAGAGCCGCC CGCTCCATCA AAGCTATCTG TAGGGGATAA AGAAGCATCT
108601 GCTCCGTAAG TTGCAATGCT CAATAGAATG GGAGTGACAA GAGTCGAAGA
108651 GATCAGGAGT TTGTGCAAGG GTATTTTCAT AGAAAGATGC TTGGGTTCAA
108701 TTAATTAACA CGTTTTCGAT AATCTAGAAA CAAAACCTAG AGCCTAGGTT
108751 TGTATTATAA TTTCGTGAAG AACTTCGTAC TTCAAAGCG AATTGACCGA
108801 AGATTTCCAT GTGGGGGTTC ACTTGAAAT GGTTCGCAGC ACGAACAGAA
108851 AAACCTTGTC GTGCGAGGTT GGTACCATAG GCCATCCAGT TAGCATCGCT
108901 AGCTATTAGG GAAGTTTGAC ATTTAGGATT GCGTCGGTAA GCATCGAGTA
108951 TATACATAAG AGTAAGATCG TAAGTTCCT TTTCTGATTT TGAGTCTCTT
109001 TCGAAGGTGA CGCCTATAGG AATCTCTACG TTGATAAGCT CGCTTTTATT
109051 GAAAGCGCGT CCTTCAGCAT GACGCTCGTA GAAGTCTTGC TGATGCGCAT
109101 AGATATACTG TACTTTGACA AAAGGTTCTGA CTTCTTTCAG AAGATACGGA
109151 ACGGAAATAA CAAAAGGCAG GCTAGCTCCA AGATCTGCAC AGAAGGCATC
109201 GTTCTCTCAA GAACCCTTGA TGATAGAGTT ATCGGTATAA TATGTCTTCA
109251 TGTGGTTGTC TGTATGGAGA TAACTGAATT TAGCATCGAA CGATAAAGGA

109301 ATGATCTGGG AGATCTCAGA GAGCACCCAG GGAGCTCGGG TTGCTTTTCC
 109351 CCAGAGGAAA TTGGCGATGT CGAAGAGCCC TTCTGTATGG TGGAAATACA
 109401 AAGAGGCACC GTAAGTATCT CCGTGGTTCT TACCTGTAAT ATGATTGCGA
 109451 TCTCTAGCAA AGAGCTGGCA GAAGGCAAAA GTAAGCTGAT CCTCGGCAGG
 109501 AGTTGTTGCT GTGATCCCTA GTGCATAACC CCCGCTGATA TGGCGGAAAC
 109551 CATGGCGGGT GGGCATAGAA TCTCTATAGA AGAAATTCGC AATTCCTGAA
 109601 AGCCATAGCT CACGCTCAAA AGGCTCCCCA CTGGACTTGG TTTCTATAAG
 109651 CTGATTGATC GAGCGTATAT CTATAAAGTT TCCCCATAAG CTATTTAGAG
 109701 GGAGATTACT TTTTCTCTCA GGACTAGGAA TGTATCCTGT ACGGGTCCAG
 109751 TTGATGCTTC CTATTTTGA GGATGTTGCA TTTGCCCAAG ACAACTGCCA
 109801 GTTTCCTTGA TACCCGTAGT GGGTTTCAGG TTCTTGAAGA GTCAGGGTAG
 109851 AAAGAGCTCC CAGAGTAATC GTTCCGTTGG CTCCTGCGGT GGTAAGTTCA
 109901 AGAAGAGGAT AGGTACTAGC ACTTTTAAAG TTATGATTCT CATAGAATGA
 109951 CCCTTCCGTG TCAATAAGCG CAATCGTTCC CGATAGGCTG ATATTTTAT
 110001 CTGCAGCTTC TGTTTTTAAA GCTGCCTTGT TGGTTCCATC TAAAGAGGAG
 110051 AGATTTACTG CTAAGCCATT AAGCGAAAGA TTTGCCTCTT TAGCACTAAG
 110101 TGTAGTCCCC CCATCCATTA AGATGCGGGA TCCTGGACTT TGAGTCAGAT
 110151 CCTTGAAAGT TACGGTGA CTATCACGAA GTACAAGATC TCCCCGCGCT
 110201 AATACTGCAG GTTGTCGGAT AGTAGAGGTG ACGTTTGCAG CGATTGCTTT
 110251 TTCTGTAGGG GAAAGCTTTT CTCCAGAAAA GACAATCGCA CCCCATACT
 110301 CGATCTCACT GTTCGCATCT GCTAAGTTTA AGTTCAATGT GTCGGTAGAA
 110351 GCTGCGGTTT CTGGATTTGT GATGGGATCA TAGAAATAGA TAGATTGCCC
 110401 CGTAGCAGCT CGTATCGATG TGACTTTAGC GGTATCAATG ATATTTATTG
 110451 CGTTTCTTGT ACTTGTGCTT CCGTTGGTGA CTTGGTTGTT ATTGAAGGTA
 110501 ATATCTCCAG AAGTAGCAGA GAGAGCGAGT TCCCCAGCAG ATGCTATATT
 110551 GATCGCTCCT CCTCCTCCCT GACCGGCGCT ACTTCCTGAG ATATTACTTT
 110601 GAAATAGAGT AGGACCTCCA GCGGAAATAC TGACCTTGAG TCCAGAGATG

110651 GCTCCGCCAT ATGTCAATGC TGTATTATTT GTGAAAGAGA GGTTTTTGT
 110701 CCCAGTAAGA GTCAGTGTGTT TATCTGTCGT AGTGCAACAA ATAGCCCCGC
 110751 CCTGAGCTTG AGCGGCTTCC CAAGCACTAT TGCCGTCAAA GATCACTTGA
 110801 AAGTTATCTG TAATCGAACA GTTGTCTAGTG CTGTACAGAG CACCGCCAGA
 110851 TCCTTTCGCT AGGTTTTGAG AGAAGGAAAC TATCCCAGGG CTGTTCTCGA
 110901 TAGTTATAGT TCCTGTAGCG TAAACTACAC CGCCTTGCTT CCCTGTGAAG
 110951 GCTTGGTTTC TCGAAAAGCT CGCAAACCTGA GATGTCCCTG ATAATAAGAA
 111001 GTTTTTTCGTA TTGATAACAC CGCCGTTATC TGACGAGAAG TTCTGAGTAA
 111051 ATATAATTTG GGAATTGCCA GTTAGAGATA GATTCCCCAC AGATTTTAAA
 111101 GCACATTGTC CAGTAGGAGA GAGAAGAAGA GAGGGACAAG AGATAATAGA
 111151 GAGTCTAGAA AAATCATTAA AGAGAAGATT CTTATCTGCT GCTGAGGTAC
 111201 TGGCTACAGT TCCAGCGCTA GAGCCCGCAT TGATAAATGC AAACCTCAGT
 111251 GCATGTTGAT TTCCTTGAA AGTAAGATCG CCGCCCGCTT CTAGGAAGCA
 111301 TCCTGAGGCT AAGGGAATTC CTAAAGCCCC TGCATTTTGA AAGGATACGT
 111351 CGGAAAGTAA GGAATAGGTA GTTCCTGCAG CAGCGTCCGT AGTGGAAGAA
 111401 ACCGTGAAGG TAGTTCGGTT AGATCCATCA TAGCTATTAT TGCTGCTATC
 111451 TAAGGTCACC TCTGCCGCGA CTATAGAGAG CGATGAAAAG AGCGGGATTG
 111501 AAGAAAAGAA CAACCAAGAG ACAGAGGACT TCATTTGTAA GCACTTTTTT
 111551 GAAACAAGGA AATTAAATTA GCAAATACTG TAAAGAAAAA AAGAAATCAA
 111601 GGGAAACGCA AGGAATTGAT TGATGCGGAG AATCAGAACC CCAAGGATGG
 111651 CGGATCTTTT ACTTCTCTTC ATACGGATCC TAAGAATCTC TTTGATGAAG
 111701 AGGGGATGCC CTCCCCCTCT GATACCCTAC AGTGCGATCT CAATAACGTA
 111751 TTCATCTTTA TAAAAAGTAT GTTTTTCTAA GATTCTCGGA GAATCTTAGA
 111801 AAGAATAACG AGTTCCACAG TTTGCATTAT AGCTTCTTGA GGAGCTGCGC
 111851 AGTTCACAAC TTCCAGAAGC GAAGCAGTCA AGACCATGAA GTAACCTCAG
 111901 ATGTCCAGAA GCCTCAGCAA AGAAAGCTTG TCGTGATAAG TTTGTAGCAA
 111951 ACGTAGACCA CGAGGTGCCA TTTGTTAAGG AGGTCAGGCA GTGAGGGTGA

112001 TCCCGGTAAG CATCTACAGC GTAACCTAAA GTAAGAAGCA AAGCACTGGG
 112051 GGGCTTTGCT GATTTCGTGTT TGAAGGTGAG TCCCATAGGG ATAGACACGT
 112101 TGACCAGATG GCTAGCGTCA AAGATACGTG GATCAGCAGC AACCTCTTGG
 112151 AATCCTTTTT GATTTTAACT CACAACCTGG AGTTTCACAT AGGGAGAGTA
 112201 GCTGGTAAGG TATCTGTAGT TTAGATCTAC AGGAAGAGAA CCACCGACTT
 112251 CAACAGCGAA GCTATGGCTG TCCCAGTCTG ATTTCCCTTG TGTGTTGTTT
 112301 GCAAGCTTTG TCGTCATATT ATGGTGGTTT CTTCCATAGG AACTTGACC
 112351 ATGGAGAACA AGGGGAGTTT CTCCTGGGAG CTCTGGAAGG ACCTTAGAGA
 112401 GGACGTGGCG ACGTAATGAG CTATGCAGGG GAATGACATA AGAGCTCTGA
 112451 GCACAGAGAG ATCCTGCATA GACTTGAGAT TTAATATCCG AGACTACGTA
 112501 ATCCTTAGAT TTGCCAAAGA GTTGGCTGAA TGCAACAGCA AAGGTATATT
 112551 CTTGAGGGGT GGTCATGCTG CCACCAACAA TATAACCTCT GGAAATCAAA
 112601 CGGAATCCTG CATTTTCCTT TTGCTTGTCT TGATGGAAGG CGTTGCCAAT
 112651 ACCTCCAATC CAAATCCCTG GATGTGAGGG AGCGTCCGAC ATCGCAGTGG
 112701 CGATCTCCTG CTGTATAGAA TGGATGTTTA CATAAGCATT CCAAAGGCTA
 112751 TTAGGAACATA AAGTCGCACG AAGCTCTGGT TTAGGAGTGT ATCCTAACGC
 112801 TTGCCATTCC GCGACCAAAG TCACCTTCCC TCCAGCTCCT ACTTTAGGAA
 112851 CCAGAGTCCA ACTCCCTTGA TACCCATAAT CCGGAGCAGC CATGCTAGAA
 112901 GGAATCGGAT TGAAGTCGTC TAAATTTACA GTTCCTGAAG TAGAAGAAAG
 112951 ATCTAAGAAA GGAAGATTTA AGTTTGCTTT CAACCCAGGA TTGTCATAGA
 113001 AACTTCCTTC ATTGTTATGG AATTTTCAGT CCCCTGAGAT TTTTAATCCC
 113051 CCACTTGTGC TGTTTACGGC AATCGTTATC ATACGCTTGC CATCTAAAGC
 113101 ATCCAGATTT ACAGAGAGAT TCTTTAGATC GATGCTGCCA TCTGTATTGT
 113151 TAGTTGTCGT GGTCTCTAAG GTCGTTCTTG CATCCATGAA TACTGTAGAA
 113201 TCAGGCTGCT GTGTGAAGGA ATATACTTGT AGGGTGGCTC CTTCTTTTAA
 113251 AACGACATTT CCTCCTGCTA AGTTGATCTT CTGGTTCAGT ATGGTGGTAG
 113301 TATTTGCAGG AATCGAGGCA TCTTGAAGTG GGAGTTTCC AGAAGAAAAT

113351 ACTATAGTTC CCGTGTTTGG GTTTCAGGT GCTACAGGGA CTACAGGCAC
 113401 TGAAGCTATA GGACCATTTT TTGGTTGGGG AGGAGGAACA ATAGCTTTGA
 113451 CAACAGGATT GATGACTAAC TCCTCTATTG TTCCTCCAGA TGCAGGAGCT
 113501 TCCATCGTAA TAGGATCATA AAAATAAATC GTATGACCAG GAGCTGCTGC
 113551 AAGCTTAGTG ATCTTAGCCC CTGCACCTAA ATGGATCGAG TTGGGAGTTG
 113601 AAGTTCCTC AGTCGCTCGG TTCCCTGAGA AAGTAATATC CCCATCAATA
 113651 GCCTCTAAGG AAAGTTCTCC GCTATCGGCT ATATAAATGG CGCCTCCCTT
 113701 GCCTCCAGAA TTATTGGTAA AGGAGACAGG ACCGTTAGCT GTAATCGAAA
 113751 GGTTTTTCGA ATAAATCGCT CCTCCCGAAG TTTCAGCAGT ATTGCCATCA
 113801 AAGTTTATGG ATTCACTGCC TGAGATTACA CACTTAGGAG CATAAATACC
 113851 ACCACCACTT CTTTTTGCCG TATTGTTAAT GAAACTTAAA CTCTCATTTT
 113901 CAGTAAGAGT TAAGCTTTTT GTAGCTATGT CAGACTCTGA GATATTACAG
 113951 AGGATCGCTC CACCACAACC TTCTTGATCT GTAGTTGTTG TTGCTGTTGC
 114001 TGTGCTGAA TTTCCAGAAA ATACAAGAGC CTTATTTTTG GTAAAGGAAG
 114051 TATTTCTTTT AGTATGTAGA GCCCCTGCTG TCTTTGCTGT ATTTGTGCTG
 114101 AAGGTCACGG TTCCCGTACT TCCCGTAAGA GTAAATCTT CGGTTTCTGT
 114151 ATAGATCGCT CCTCCTGCAG TTTCGGCAGT ATTGCCATCA AAGGTAAGAG
 114201 TCGTGTTTCC ATGCAGAGCA CACTTGGTCG CATAGATCGC ACCGCCACTT
 114251 ACTGTTGCAG TATTACCAGA GAGACTCAG TTTTCGTTAT CTTCAATCCA
 114301 GAGTCCTTTT TTAGTACTTA CAGATGCTGA CTCAAGAAAC GATAGGATTG
 114351 CCCACCGCA ACCCTCTGA TTTGCTGAAG AATTACTCGG GCCCGTAGCT
 114401 TTGTTCCCTG AAAAGAGCAG GTTGGTATTA CCAGACAGAG AGTTGTTGCC
 114451 TTTAGAATAT AAGGCGCCGC CTGTCTTTGC TGTATTTGTG CTGAAGGTCA
 114501 CGGTTCTGT ACTTCCTGTA AGAGTAAAT CTTAGTTTC TGTATAGATC
 114551 GCCCCTCCTG AAGTTCAGC AGTATTGCCG TCAAAGGTCA GGGAGCCGTT
 114601 TCCAGTTAGA GTACATTTGG TAGCATAGAT CGCACCACCA CTTACTGTTG
 114651 CAGCATTACT AGTGAGGCTG ACTTCTTGGT TGTTCGAAT CGATAGTCCT

114701 GTTTTATCGC TTACGGATCC TGAATCAATA AAGGCTAGGA TTGCCCCACC
 114751 GCAACCCTCT TGATTGCTG AAGAATTACT CGGGCCCGTA GCTTTGTTCC
 114801 CTGAAAAGAG CAGGTTGGTA TTTCCAGTCA GCGAGCTGTT TCCTTTAGAA
 114851 TATAAGGCGC CGCCTGTCTT TGCTGTATTT GTGCTGAAGG TCACGGTTCC
 114901 CGTACTTCCC TTAAGAGAAA AATCTTCAGT TTCTGTATAG ATAGCTCCGC
 114951 CACATCCTGC TGTCGCAGTA TTCTGATCGA AGGTAAGAGT TGTGTTTCCA
 115001 TCCAGAGTAC ATTTAGTAGC GTAGATCGCT CCACCATTCTG CAGTTGTTGT
 115051 ATTACTAGTG AAGCTCATTT CTGATTCTG AGAAATGGCT AATCCAGTTT
 115101 TGTCTGTTGC TGTAGCAAGA TAACAACAGA TTGCCCCACC ACAACCTTCC
 115151 GGGTTATTTG CCTGTGCTGC TGAGCCGGTT GTTTTATTTT CCTGAAAAG
 115201 TACTTGAGTG TTGCCGGTAA GAGCAAGATT GTCATCAGAG CTCCAAGCAC
 115251 CCCCCGTCTT TGCAGTATTA GATTGAAGG TAACGACTCC TGTATTGGCA
 115301 TCTAGCGTGC TATCCTTTTC TTTGAGTAG ATCCCCCAC CTTTATCTGT
 115351 AGCAGTATTT GAGGAGAAGG TCACCGTTCC TGAGTTTCCT TGGACTGTAG
 115401 TGTTTGCTGT ACTACAGAGG GCCCCGCCAT TTTTGTGCT AGTATTTTGA
 115451 TCTAAGAGAG CTGCTGTCGT AGTCTTAGCA AGATCGATGC TGTAGGCAGA
 115501 AACTGCAGCT CCATCTTTTT CTGAAGTATT TTTTGGAGG GTGACACTGG
 115551 CATTGTCAGT AAAAGTCGCA GTACCTCCCT CTGTATTTGT CACACAAATA
 115601 GCACCCTTGC CGCCCGAAGT TCCTGTTGCT GGAGCTGAGT CGATTAAGAG
 115651 TGACGAGAAT CCTGAGAAAG AAAGAGCTGT GTTGGTATTG TTAATTGCAG
 115701 CACCATCATG CGTAAGCGCT ATGGTTTGCA GAACCAATGA GTGATCAGCT
 115751 CCAACAAAAC TCAATGCTCC TCCTGTGTTT GTAAAACAGC TTTTATCTGC
 115801 AGGAGTAATT GCAGATACAT TCGTAATAGA AACATCGCTA GTGAGAGTGT
 115851 AGGTAGTTCC TGAAGCATCC GAAGTTTCCT TGGCAGTGAA TGCTGCGCTA
 115901 CCACTACTAC CATTTTCATA GTTATCGGAT GATGAGAGAT CCGTGTTAGC
 115951 AGCCATTAGT GGATGTAGGG AGAAAAC^TTAA AGCCGAAGAG GTAAGTAGCC
 116001 AAGGTAAAGA ATATTT^TCATG TGTCTTTGGG GAAAAGCTTT TTATCAAAAA

116051 TACTCCCATA GCATGTGGCT TTAGGAGCAT GGTGCACCAA TAGAGAATAC
116101 AGTTAAATAA ATCAAGTAAA TGCTCTGGAG AAGACTCTCA GTTATAGAAG
116151 TTTCAATCTT GGGAGAGAAG CATTTAAGGT ATTTTCTAT ATTTAAGAGT
116201 CCCTTAAAC ATAAGGAAA TGCTTAAGGG TAGGGGAGAA GGTGTACAAG
116251 CGGTTTTGCT TTTAGACCTT CTTGAATTTT AGAAGGAGAG AGTAAGGAAG
116301 ATGGTTATCT AAACCACGGA TCCTTATTG TTCTTCTGT GTTCCGTTT
116351 TTCTCTTCG TCATAACCGT CAGAGAATGG ATTGGAGGGG CTGAGGTTAG
116401 CCGTGCTTTT TCCGTCTGGA GGCGGAGTTT TAAGAGATTG ATTCGATTTT
116451 CCGCTTTCT TCTTAGATTG CTGTTTCTGT TCTTCATCTT GATTTTGCTG
116501 TTGCTGCTGT TTATCTTGGC GATCACGAGG GGAACGGCGG GAAGATGAAG
116551 AAACGAGG AATTGAGGGT TCTTTCCTC CCTTAGGGTA GACCGGCTCA
116601 GGGTGTAGAA CCGTCCCGT GGAGAAATTA GGAGAGCGGG AACTTGCAGG
116651 CATTATGGGT GTAAACGCAC TGCTCGCTCC ACTACCGAAT GAACTGCTTC
116701 TTAAATCTTT GAAGTGATAA GGCTGAAAAT CGTCCTTAAA TGGAGGACTT
116751 ATAGATGCGG GTCGTGTAGA AGCCGCATCT GAGGGTTTTG TTTTAAACG
116801 TGAGAGGCTG CCGAAGGAGC GTCTATGCTC AGGATTTCGT GACGAATAGA
116851 AGAAAGATCC TTGAGCCTCC CGACCGATCC TACGATGACG GGCATCTTGG
116901 CGTTGTGAAG CTTCGTGTTT CTCCATGCGG GCAGATGCAG AAAGATCTTT
116951 TTTCTTTTCC ACTGCGATTT TTTTGCATC GTCGGCGGAT GGAGTCACTA
117001 GAATAGGCTC CGCGGTCTCT TGGGTTTTCT TTTCCATCCT CAAGAGGTGT
117051 TCCCTACGGA AATAATCTAA GGTGAGCTTA TTCAAGGACA TGAAGTATAG
117101 CCCCACCGTA ATCAATCCCA TGACAAACAT AGGGTTGGCA AAGACTAACA
117151 TCGTCCCACT AGAGGCAACG AAAGCCCCTG CAATAAGACC CGCAGCAATC
117201 GCTAATACAA TGATAGGAAC GACGATAGCA GTGATTGCCT CACCGATTTT
117251 CTTGGCCTTC GGAAGTGTCA GAATATCAGA AATAAGTAGA GTAAGTCCCA
117301 AAGCTCCCAG GGCAAGTGCC GGAGCGAGAG CATAAAGCAT AAGGCTGTTT
117351 CCTGAAGCAG TCAGGAGAAT CGAAAGGATC GCAATTGCCG CAAGGGCAAT

117401 AATGCCTGTG TCATAGACAT AGCGCATTTT AGGATATTTA TCAGGAATAG
 117451 TAATAAATCT TTTGAATAAC CCTGTGGATG ACGTTTTTAA ACGTTTTTACC
 117501 ACGCTTGGCT GCCCTGATGG GGATGCTGCT GCAGGAACTT GAGGTTGACC
 117551 TAAAGGGTTT ATAGGGGGTT GGCTCATACT ATCAACTTAC TGTAATTATC
 117601 ATTAGGCCCCA TGAATTTTCA TTCATAGGAT ATATTTTATA CTATTATAAG
 117651 ATTTAATAGG ATTTAGTTAG TTCTCTTTTC TTCTGAGTCT TAACTTTTTT
 117701 ATTAAATAAA GTTTATTTGT TAAAATCTTA ACAGATTTTT AACTAAAAC
 117751 TTAAGTTATT TTTATTTGGA ACTTTTAGTC GAAATAAGAC TCGCTTATGA
 117801 GAGGGACATA CTCATCAGCA AATGGAGGGG GCGTGTGAGG TCGTGAGGGT
 117851 GGAGGCTCAT ACGGGGGGAA AAGATTTTTA TGCACTTGGA AGAGGGTAAC
 117901 GCTAGTAGTT ACAGACATGA GAAGCCCCAA GGAGATAAAG ACAACTGGAG
 117951 GGGGGATAAA GACCAGACTA AAGACCAGAC CGATAATCAC AGCAATCGTA
 118001 AGAATGTGGA GGAGCCAAGC CATAGCATAC GAAGCAAGAT GTTGGCAGCT
 118051 TTTTGTTTTT ATAGCGTTAA AGAGAGCCCC TAGGGGTAGG GTCAAGGCAG
 118101 CATATGTAGA AGCGCAAGGG ATCAGGATAA CCTTAAGAGC TCCTAAGACC
 118151 GAGGATACCA GTACTTCTAT GGTAAGAGCA GTTCTGGGAT AGCTCTTCGC
 118201 ACAGTTTTTT AATTTTCGAG CTAGTATTCG TTCCGGAAAA ATGCCATTCA
 118251 GGTATAGCTG AGAGCCTTGT TTGCAGATAT TTTTGAATCC CATATCCGTT
 118301 TGAAAGAGAA TATTTTATGA AAAATTATGT AAAAATTCTA AGAGGATAGT
 118351 GGTTTTTAGA CAATCGAAAT TCCTGAAAAG GCAGGAAAAT GAGAGACACA
 118401 AGTAGACAAA ATCTCCTGAA GTTTTTGTAT GGGCCTGTAA AAAAATCTTT
 118451 CTGGAAACTG GAAATTAGAA GTTCATTACA GCGGAACCAC CGAAAAATGC
 118501 GGAGGTTTTA AAAGTATGGG ACGTTTTTATT ATTGTTGTGA TCATCAGCTA
 118551 TCGAGATATC ATTCCCAATA GACCAACCGT AAAATCCCTT AATAAACTT
 118601 CCCGGCCAAG GGGTGAGCTT CACGTTAGCT TCGATTTCAC GTCCTTGATA
 118651 TTCAAAGATG CCGCGAGAAG AAATTAGGTG ATTTTTGTGA AGTTTTTTTC
 118701 GGAATCTTGT AAGGCGGTAA GCAAGCCCTA AGTTACAGAC AGATGTCGAG

118751 CGGTAATCAA TAGAAAAATT CACAGGATAG ATGCAATTGA GAGTTAGTTG
 118801 GTCGGTAGCC TTGTAATAA CACCTACTAA AGGCCAAGCC TTCTCTTGAT
 118851 GGAGGCCTGT TTCATTAATG ACGCCAAAAA TAGCAGAAAAG CTTCTCAGTG
 118901 GCCTGGTATT TTCCAGAAAAG AACTCCTTGA TAGAGTCCAT AACCCATCTC
 118951 AATATTTTTTA GGATCCACAA GCCCAGAAAAG AATGATAGAC CACTGCCAAT
 119001 TTTTAAAGGA GAGTGTATAA GCTCCTAAAG AGAGGAGAAC ATAGTTATAA
 119051 AAAGAAGTAT CTTGGAAAGT CGCCCACCCA AGTCCATTAG GATCTGTCTC
 119101 CGAAATAGGA AGTGAGCTTT TCCATTGAAT ATCCGCACCT ATATAGCCAG
 119151 TAGAAAACAG TAGCCCAGAA TGCTCTGTAA TCGGAAGTGT GCAGAGAAAC
 119201 GTTCCATCGT ATTGACGATA GCCTATAGTT TGATGAGGCA GCTTTTTTAA
 119251 TTTAGCATCG TTCACCTTTA GGTATTGTAC CTGAGCAGAG AAAGGACGTG
 119301 GAGGAGGATT TTTACATGCT TCTTCATCAA TTCCACAAGC ATCTTGAACA
 119351 ATAAAAATAG GAGTCGAGAG TACGTGTCCC GCAAATGCAG CGATGTGGAA
 119401 GAGCAGTTTG AACATGTTCT GTAAGATTCT CCAACGTTAC TAGAGATTGA
 119451 AATGGAATAT ACGTAATTTT TAAATTACTA TCTAATAATT TTCCTACTCA
 119501 GGAGGGACTT TCAAGAAGAT TTCCCTAGAT TTAGGGGGAT GAAAGGACTA
 119551 GAATTTTTCT AATCAGCAAT CGAACAAAAT TACAGGGTTC TTGAGAGGT
 119601 GCTTAATACT ACGCATAAAT GATTTCTGAG ATGATTTCCG TTTCGGCTGT
 119651 TTCTGTGCTA ACTTGTTTTG GCACTCAGTA ATTTTAAGCT CAAGATTTTG
 119701 GATCGATTCC TTTTTCGCTT CAATTCATT TAAAATTCTT GTTTTTTCTT
 119751 GATTTTGTGT TGTCAGCTCT GGAAGATGCA CGCTTTTAAT TGTAGAGAAC
 119801 ATCAACCCGT ATCTCTGAAC AAATCCTGAT CCTATAGAGG AAATGTCTTG
 119851 AGATATCTCA TCAATCAGTT TTGTTCTGT CTGCTTATTT TTAAGAGGGA
 119901 TTGCTAGAAG AAGAGCAAGA ATCAGTGTGA GTACGATAAT TCCAAGCCG
 119951 ATGCCACAGA TGATCCAGTT TCCAGTATAT CCTGCGTAAC CTGCTGAGAT
 120001 CGTCCCCCT AAGGTTAACA GTAGTGTTAA GGCAAAACAG AGCGTATGCT
 120051 TTGTGGAAAG TTTGGAGCGA AAGGTCTCCC AAGGAGCCGC TGAATCGGA

120101 TGAGGCTGCG TAATATAAGC ATTAGGAAGA TTGAGAACTT CTGTAGCATG
 120151 AGAGAGAGGG CTGCTCTCGG GGAAGTGGTGA TGGGGCTACG GAAGTTGCCA
 120201 TGTGTTTCC TCGAATCGTT GCTAACTAGT TTTGATTGT CTTTTCATTC
 120251 TTGTGAGAAA GCTCAAAACG TTTTGTGATAA AGAGTTGATT GCAGTTTGAA
 120301 CTCTTTTGT GAGAGGTCGC AGAGTTCTTC GTATAACTTT AGCGTATCTT
 120351 GAGTGGTTTT TCTCTGTGCA TTGGTTACAA GCTGTAGAGA GGATTCAGGT
 120401 CGAATCTTCT CCCTGATAAA TTGTACTGTC GTATGGAGCT GCTGAGGGAG
 120451 CTGCCGGATG AATTTAATAA ATCCTACCAA GGCTTGTAGG CAGAGAAGAG
 120501 TCAAAATAGT AAGAACAATG CCAACGGCAA TCAACAGAAT GCTTTGGCTA
 120551 TAGCAACCCA AACAGATAAT AGCGATTCCA GCAAGAGCTA CAATCACTAA
 120601 GATCGTAATC GCAGCAATAT GCATAGGAAT GGAGTGGGTG AGAAAAACAT
 120651 TGCCCTTCTC CCCCAAAGCT ACGATCTCCT TATTCGTAAT GAATAAATCA
 120701 GCAGACTCTT CCGGAAGGGA TGAGGGAAAT ACCCCGTTTA AAGTACTAGA
 120751 CACAAAGAGA ACTCTATTAT TTGAGGAAAT AATTTAAGAA AAATGGTATT
 120801 TTTAGTCAAT TAGTAAGCGA GTCATGCCTC TTAGTTATTC AAATTTTAA
 120851 AACCTTACCC TTCCTATGAG GAGACAAGTA AGAGAAATTA TGCAACAAAC
 120901 TGTAATTGTA GCAATGTCAG GAGGCGTGGA TTCTTCTGTC GTTGCCTATT
 120951 TATTCAAAAA ATTTACCAAT TATAAGGTTA TTGGCATCTT CATGAAGAAT
 121001 TGGGAAGAGG ATCGCGACGG CGGTCTCAGC TCGACTACTA AAGATTATGA
 121051 TGATGTCGAG AGGGTCTGTC TTCAGCTCGA TATACAGTAT TACACCGTAT
 121101 CTTTTGCTAA AGAATATAGA GAAAGAGTGT TCGCTCGTTT CCTCAAGGAA
 121151 TACTCTTTAG GCTACACTCC TAACCCCGAC ATTCTTTGTA ACCGAGAAAT
 121201 CAAATTTGAC CTTCTACAAA AGAAAGTCCA GGAAGTTGGC GGAGATTACC
 121251 TCGCTACAGG GCACTACTGC CGATTAAATA CCGAGCTCCA AGAAACCCAA
 121301 CTCCTTAGAG GTTGCGATCC TCAAAAAGAT CAGAGCTATT TTTTATCAGG
 121351 AACTCCTAAA AGTGCTCTTC ACAATGTGCT CTTTCCTCTT GGGGAAATGA
 121401 ATAAGACTGA AGTTCGTGCG ATTGCAGCTC AAGCAGCTCT TCCCACAGCA

121451 GAAAAAAAG ATAGTACAGG CATTTGCTTT ATAGGGAAGC GCCCTTTTAA
 121501 AGAGTTCCTA GAGAAGTTTC TTCCCAATAA AACAGGCAAC GTTATCGATT
 121551 GGGATACCAA GGAAATTGTA GGGCAACATC AGGGAGCTCA CTATTATACT
 121601 ATAGGGCAGC GGCGAGGACT TGATCTTGA GGATCCGAGA AACCTGTTA
 121651 TGTGTGGGA AAAAATATAG AGGAAAATAG CATTTATATT GTGAGGGGGG
 121701 AAGACCATCC CCAGCTCTAC CTACGGGAAT TAACAGCTAG AGAGCTCAAT
 121751 TGGTTTACCC CTCCTAAATC CGGATGTCAC TGTAGCGCTA AAGTCCGCTA
 121801 CCGTTCTCCT GATGAAGCTT GCACGATAGA TTATAGCTCA GGTGACGAGG
 121851 TCAAGGTGCG ATTTTCACAA CCCGTCAAGG CGGTAACCTC AGGACAAACA
 121901 ATAGCGTTTT ATCAAGGAGA TACCTGCCTT GGTAGTGGAG TTATCGACGT
 121951 TCCTATGATT CCAAGTGAGG GCTAGGGAGA GCAGCTTCCT GCTCCTCTTC
 122001 TTCCCTTTCA AAGGCAACGC GATTTTCAAC CAAGGTTGCT CGTAGCTTGC
 122051 GAGCTTCTTG ACGGCAGGAC TCTTTAAGCA AGAGCTCCGC TAGAGGATCT
 122101 TCAAGGTA CTCAATGAC ACGGCGTAGA GGACGTGCTC CCATTTCTGG
 122151 AGAATGCCCC TTCGTTACTA GGAAGGAAAT CACAGAGTCT GGGATGTTCA
 122201 AAGCCATTTG GTAGTTTTTC AGTCTCGAGT CCAGTTTGTT GATCTCTAAA
 122251 TGGATGATCT CCGATAGAGA TTCTTTCTCG AGGGGACGGA AAATCACACT
 122301 TTCATCCAAA CGGTTAATGA ACTCAGGCTT TAAGTGTTTC TTCATAGCAT
 122351 GTTCGATTTT CTCTTGGATG ACCTTATAGT CCATATGGGA CTCAAGCCA
 122401 AAACCAATTT CTCCGCTTTT ACGAATGAGA TCAGCTCCCA AATTGGAGGT
 122451 CATGATAATA ATGGCATGAC GGAAATCCAC TTTGCGACCA AAAGAATCAG
 122501 TAAGACGTCC TTGCTCTAAA ATTTGCAACA TCAGGTCCAT AATGTCTGGG
 122551 TGTGCCTTTT CTATCTCATC AAAGAGAACA ACGCAGTAAG GACGGCGACG
 122601 TACCTGTTCC GTAAGGTGGC CCCCTTCTTC ATGACCTACA TATCCTGGAG
 122651 GTGATCCCAT CATCTTGGTA GCAGCAAATT TCTCCATGTA CTCTGACATG
 122701 TCTACCTGAA TCAGAGCGTC TTCACCACCG AACATCTCTA TAGCAATTTG
 122751 TTGGGCGAGC AGGCTTTTCC CTACACCGGT AGGCCCAAGG AATAGGAAGG

122801 AGCCCGTAGG TCGGTTAGGA TCTTTGATCC CTGTTGAGA ACGTCGGATG
 122851 GCACGGCAAA TGCTGGTAAC GGCATCATTT TGACCAATGA CTTTCTTCT
 122901 TAACGTGTCT TCTAACTTCA GAAGCTTCTC ACTTTCAGCT TCTGTGAGCC
 122951 TTGCTGAGGG AATTCCTGTT TGTAGAGAAA CTACCTGAGC GACTGCTTCT
 123001 TCATCTACAG GAACTTGGTG CTCTTCTTTA TGATTTTCCC ATTCTGTCT
 123051 CATACTTTGC AGACGTTCGC GAAGTTTTTT CTCTTCATCA CGTAAACCTG
 123101 CAGCTTTTTT GTATTCTTGA GTTCCAATGG CCTGCTCTTT GGCCAATTTT
 123151 GTATTTTCGA TTTCAGCCTC TAGCTTCATT AAATCTGTAG GCTGACCCAT
 123201 TGTATTCACA CGGACACGAG CCCCAGCTTC ATCTAAAAGA TCTATTGCTT
 123251 TATCAGGGAG GAAACGTCCA TGAACATATT GATCAGAAAG AGTCGCAGCT
 123301 GCTTTTAAAG CTTCTTCAGT AATGAAGACA TTGTGATGTT CTTCACTT
 123351 TTTCTTGAGG CCACGTAAAA TCTCAATAGT CTCATCTACA CTAGGAGGGT
 123401 GAACCACGAT TTTTGGAAA CGACGTCTA AAGCTGCGTC TTTTCTATG
 123451 TGCTTGCGAT ACTCATCTAT CGTAGTTGCT CCAATACACT GAATTCACC
 123501 TCGCGCTAAC GCAGGTTTTA AAATGTTTGA AGCATCGATA GCACCTTCAG
 123551 CTGCTCCTGC TCCTACAATC GTGTGGAGCT CGTCAATGAA GAGCAAGATG
 123601 TTTCCATGCT TGCGAACTTC ATCCATGACA GCTTTGATCC GTTCCTCAA
 123651 TTGCCCTCGA TATTTGTTC CAGCAATCAT TAATGCTAGA TCTAGAGTAA
 123701 TCAGTCGCTT TTTCCGTAAG GCATCAGGAA CCTCATTCAG AATGATTTTT
 123751 TGAGCCAGAC CCTCAACAAT TGCAGTCTTA CCAACTCCAG CTCTCCAAT
 123801 AAGTACAGGA TTGTTTTTTC TTCTTCGGCA AAGAATCAAA ATCAACCGTT
 123851 CGACTTCTGA AGAACGACCA ATGACAGGAT CGAGCTTAGA CTCTCGGACC
 123901 ATCTCCGTTA AATCATAACC ATATGCTTTC AGAGCAGAAA GCTTTTCGTT
 123951 TTTGTCAGAA CTAAGCTAT GACCTAAAGG AGATTTTGAA GATGAAGGGT
 124001 TGCTTCGAGA GGATGAGGAA GAAGACGACG ACGAAGGAGG AAGTTGTAGA
 124051 TTGAAGGTCT CTAATTCTCT AAGAATTTC TTACGAACCT CTCTTGATC
 124101 GATATGTAAG TTTTCTAATA CCTGAAGAGC GACACTATCT GATTGATGTA

124151 GGATCCCTAA GAGTAAATGC TCCGTCCCGA CATAATTGTG CTCTAAAAGG
 124201 CTGGCCTCTT CATTTGCTGA TTCAAAGAT TTTTCTACTC TTCCTGTAAG
 124251 GGCAGGGTCT CCGTAGACTT GAATTTCTGG ACCATAACCA ATCAGGCGTT
 124301 CCACCTCTTG CCGTGCCGTA TCAAAATCTA TACCGAGGTT GCGTAATACA
 124351 TTAACAGCTA CCCCTTGACC AAGTTTGAGA AGACCAAGCA GGATGTGCTC
 124401 AGTACCCAGG TAGTTATGAT TTAAACGCTG AGCCTCCTTT TTCGCCAGTT
 124451 TAATGACTTG TTTTGCTCTA TTAGTGAAC TCTCAAACAT AAAAACCTAA
 124501 AAGACAGGGG TAGAACTTTC CTTAAGCATA TACGAAATTT AAAATAATGA
 124551 TGCAACTCTT CGCTCTAAAC CAGCAAATTT GGTAATAATC CTCTGAGTTT
 124601 AAGGGAAAGT TATGCACAAA CCTTTTGTAT ATGATACAAT AGTTCAGCTT
 124651 CTTTTGAAAC AGTCTTAATT AGTTTATGT TTGTTATATG AAAGTTCGTA
 124701 TCGTAGATTC AGGAAAATCT TCAGCGGCCT CCCACATGGC TAAGGACAGA
 124751 GATTTATTAG AATCTCTGCA AGATGGGGAG CTCATTTTAC ACCTTTATGA
 124801 GTGGGAGAAT CCTTGTTCTC TGACGTACGG TCACTTTATG CGTCCAGAAA
 124851 AATTTTTACT TTCCAATAT GCGGATCTAG GATTGGACGC CGCAGTGCGG
 124901 CCTACGGGAG GGGGATTTGT CTTCCATAAG GGAGATTATG CTTTTTCTGT
 124951 TCTTATGTCT GCGACACATC CTTCTATTTC TTCTTCGGTA CTGAGAACT
 125001 ACCATACTGT AACTCTTTT GTAGCGAAGG TTCTAGAGAA AGTATTTTCGG
 125051 ATCCAGGGAA TGTTAGCTCC AGAAGACGAA AACTCTTCTT CCAGAGATTC
 125101 AGGAAATTTT TGTATGGCAA AACTTCGAA GTATGACGTT CTTTTTGGGG
 125151 ACAAGAAGAT AGGGGGCGCT GCCCAACGCA AGGTGCAACA GGGATTTTAA
 125201 CATCAAGGAT CCTTATTCTT ATCGGGAAGT TCTTCTGAGT TTTACCAGAG
 125251 ATTTTAAAAA CCCGAGGTTT TTGAAGAAAT TATTGAACAA ATCCAGATTC
 125301 ACGCGTTTTT CCCTTTAGGT TTGAAGCTG CTGATGAAGT GCTGCAGGAG
 125351 GCGCGTCAGC AAGTCAAAGA GCGGTTTATT AAATTGTTTT GTGGTGAGGG
 125401 GTTATGATGA GTCGGTTGCG TTTTCGCTTG GCAGCTCTTG GAATATTTTT
 125451 TATTTTGCTG GTTCCTAATT CTGTTTCAGC AAAGACAATC GTAGCTTCAG

125501 ACAAGGAGAA GGTGGAGTT CTTGTTTATG ACAATAGTGT AGAGGCCTTT
 125551 CAACAGATAT TGGATTGCAT AGATCATGCA AATTTTTATG TAGAACTGTG
 125601 TCCCTGCATG ACAGGAGGCC GAACGCTTAA AGAGATGGTA GATCACCTCG
 125651 AGGCTCGTAT GGATCTGGTT CCAGAGCTCT GTAGCTATAT CATTATCCAA
 125701 CCCACGTTTA CCGATGCTGA AGACCAAAAA TTA CTCAAAG CTCTCAAAGA
 125751 ACGTCATCCC AACC GGTTTT TCTACGTTTT TACAGGGTGC CCACCCTCAA
 125801 CAAGCATCCT CGCTCCTAAT GTCATTGAAA TGCATATCAA ACTTTCTATC
 125851 ATCGATGGGA AATATTGTAT TTTAGGTGGT ACCAATTTTG AAGAGTTTAT
 125901 GTGCACTCCA GGGGATGAGG TTCCTGAGAA AGTGGATAAC CCACGTTTAT
 125951 TTGTCAGTGG AGTGCGTCGG CCCCTAGCAT TTCGTGATCA GGATATCATG
 126001 TTGCGTTCTA CAGCATTCGG TTTGCAGCTC AGAGAAGAAT ATCATAAGCA
 126051 ATTTGCTATG TGGGACTACT ATGCACATCA TATGTGGTTC ATTGATAATC
 126101 CTGAACAGTT TGCAGGCGCC TGCTCTCCAC TGACTTTAGA ACAAGCCGAG
 126151 GAGACAGTAT TTCCTGGATT TGACAAACAT GAAGATCTTG TTCTTGTCGA
 126201 CTCTTCCAAG ATCAGGATAG TTTTAGGTGG TCCCCACGAT AAGCAACCCA
 126251 ATCCTGTGAC TCAAGAATAT TTGAACTTA TCCAGGGAGC TAGATCTTCT
 126301 GTGAAGCTTG CTCACATGTA TTTTCATCCCT AAGGACGAGC TTTTAAATGC
 126351 TCTTGTCGAC GTTTCTCATA ATCACGGTGT TCATCTGAGT TTAATTACGA
 126401 ACGGCTGTCA TGAATTAAGT CCTGCAATTA CAGGACCCTA TGCTTGGGGA
 126451 AACCGTATTA ACTATTTTCGC CTTGCTCTAT GGGAAACGGT ATCCTCTTTG
 126501 GAAAAAATGG TTTTGCGAAA AGCTAAAACC TTATGAGCGG GTTTCTATTT
 126551 ATGAGTTTGC TATTTGGGAA ACGCAGTTGC ACAAGAAGTG TATGATTATC
 126601 GATGATGAAA TTTTGTGAT CGGAAGTTAT AATTTTGGAA AGAAAAGTGA
 126651 TGCCTTTGAT TACGAAAGTA TTGTAGTTAT CGAATCTCCA GAAGTCGCTG
 126701 CAAAAGCTAA CAAAGTCTTC AATAAAGATA TCGGATTGTC GATTCTGTGA
 126751 AGTCATGGCG ACATTTTCTC TTGGTATTTC CATTCCGTAC ACCACACTTT
 126801 GGGACATTTG CAGCTGACCT ATATGCCAGC CTAGCGTCCC TGGGTGCGAA

126851 TCTACCAACA GGATCTCTTC TGCAGGCTCT GCAGGGATCC TGCCTGGTTT
 126901 TTTTCTCTGC TATCGTTTAC ACTACGCTTT TATTGTTTGG GTAGAGGGTG
 126951 GACCTTGTTA TCGTTCTTCT ATAAGCATCA AAAAAAATTT ATCGGCATTG
 127001 TCATTGCTGT AGTTTGTGTT TCTTGGTATT GGAGTGGGT TGGGACGATT
 127051 CTCTAGAAAA GGTTCTGCAG AGTCCACCTC ACGTCGGACT GTTTTTACTA
 127101 CCGCTTCAGG GAAGCGGTAT GTAGAGAAAG ATTCATGGC TATGAAGAAG
 127151 TTCTTTGCTC ACGAAGCGTA TCCATTTACA GGAACCCTA GAGCTTGGA
 127201 TTTTATCAAT GAGGGGCTAC TTA CTGATTA TTTTCTAACG ACAAGGGTGG
 127251 GAGAAAAACT CTTTTTAAAA GTGTACCATC CGGGAGAGAA AATTTT TAGT
 127301 AAGGAGAAAG CTTACCAGCC GTATCGTCGT TTTGACGCTC CTTTTATTTC
 127351 CTCTGAAGAA GTTTGGAAAT CTT CAGCTCC CCAGCTTTTA GAGATCCTGA
 127401 AGGTCTTTCA ACAAATCGAG ACCCCATAT CAAAAGAAGG ATTTCTTGCT
 127451 AGAGCCAAGC TCTTTT TAGA AGAGAGAAGG TTCCCTCATT ATGTGCTTCG
 127501 ACAAATGTTG GAGTACCGCA GGCAAATGTT TGCTCTTCCC CCAGATGAAG
 127551 CCTTATCTCG CGGGAAAGAC TTGCGGTTAT TTGGCTACCA GACGATTCAA
 127601 GACTGGTTTG GGGATGCCTA CCTTTCTGCT GCTGTTGAGC TCTTGATCCG
 127651 CTTTATTGAC GAGCAGAAAA AAGTACTTCC CAGGCCCTCA AAACAAGAAG
 127701 CTCGTGACGA CTTTTATGAT AAGGCGAAGC ATGCCTATAC TAAGATCAGT
 127751 AAGAATAAGG AATTTTCCTT AGGATTTGAA GAATTTGTAA ACTCGTATTT
 127801 TCAGTTTTTA GAGATCTCTG AGTCCGAATT TTTCAATATG TATCGAGACA
 127851 TATTGTTGTG CAAAAGAGCT CTTCTCCTAT TGCAGGGAGG CGTTTCTTTT
 127901 GACTTCCAAC CTCTAACTAC ATTTTTCGTT CAAGGAAAAG ATTCCATACA
 127951 AGTAGAGTTC TTTAGACTCC CTAAGGAGTA TAGCTTTAAA ACAAACAAG
 128001 AGTTAAAAGC TTTCGAAGTC TATTTAAAGT TAGTGAGTTT ACCTAAATCG
 128051 GATAGTTTGG ATGTTCTTAA TGAGATCCTT CCTATAGCGA CCATAAAAGC
 128101 TAAAGAGCCT CGGT TAGTAG GCAGACGGTT TTCTATAGAC TATAAGAGAG
 128151 TCGCTTTGCA AGACTTAGCA GCTACTGTAC CTATGGTTGA AGTGCTGCAC

128201 TGGCAACAAA ATTCTGAGCA CTTCCAGGAG ATTCTCCAGC AGTTTCCTGA
 128251 CGTTGAGACG TGTCAGTCGT ATAAAGACTT CCAACATCTT AAGCCTGCGC
 128301 TGCGAGATAA AATTTCTCTT TTCACACGCA AGGAAATCTT AAGGGCCCCG
 128351 CCTGAGAGAA TTCTGCAATC GCTACAGCAA GTTCCTAAGC AGAGCCAAGA
 128401 AGTTCTCTTA TCTGCAGGGA AGAATAGTGC TCTACCAGGA ATATCCGACG
 128451 GTCAGCAATT AGCCAAAGTG TTGCTTGAAA ACGAGGTTTT AGATTTATAT
 128501 AGCCAGGATG CAGAGACCTA TTATACTATT ATTGTTAATA GTTCTTTTGA
 128551 AAAAGAAGAA GTGCTTCCTT ATCGTGAGGT TTAAAGAGA GATTTGGCCT
 128601 CACAGTTACT TACTTCTCAT GGTTCATCTG TTGACATGGA GCGTCTAGAA
 128651 TCTGCGTTGC GTACACGGTA TCCAGGAGAA GAAGGCGCTA GCCTATGGCA
 128701 ACGACGTCTT TGGAAGGTAG TGGAAAACCA CAGATTGGGA AGGCATCTCG
 128751 AGGGGTCTTT CTCTTGAGC TTAGATCGCT CATTGAAGAC TTTTCCCGA
 128801 GGAGACAAGG AGCTGCCCCA AGAGTTTGAT AGGATTTTCT CTATGAAGGT
 128851 AGGAGACTAT TCTTCTGTAT TCATGAGTCC TAACGAAGGG CCCTGTTATT
 128901 ATCAATGCCT CTCTCATTTA CTGTATGATC GTCCTGCTAG CGTGGATAAA
 128951 CTATTTTTAG CTAAAAGTCA GCTAGATGAA GAACTTTTAG GATCCTATAT
 129001 GGAACGCTTT ATAGAACAGG GAGTCGTAAG GTGATGTGGT ATTCTGATTA
 129051 TCATGTTTGG ATTTTGCCCC TCCATGAGAG GGTGGTGCGC CTCGGGTAA
 129101 CAGAAAAAAT GCAGAAAAAT TTAGGAGCCA TTCTCCATGT GGATTTACCT
 129151 TCAGTAGGGA GTCTATGTAA AGAAGGTGAG GTTTTAGTCA TTCTGGAATC
 129201 TTCTAAATCT GCTATAGAGG TGTTAAGTCC TGTATCAGGA GAGGTTATCG
 129251 ATATCAACCT TGATTTAGTG GATAATCCTC AGAAGATTAA CGAAGCTCCA
 129301 GAAGGTGAGG GATGGTTGGC TGTAGTCCGA CTAGACCAGG ACTGGGATCC
 129351 TTCTAATCTT TCTTTGATGG ATGAAGAGTA AATTTTTTAT TAGATATACT
 129401 CATTTTTTTC AGAAGATAAG AGGTATTTTT TTAAGGCTAA AACATTTAAA
 129451 ATTTATGTCT AAGGTTTAAA AAATACATCA GAATTATTCT ATGGATCCAG
 129501 CTAGTCCGGT AGCCCCTCAT GTCCTACAAG ATCATGTGCA ACTATCTTCT

129551 GAAGAATTGT CCGCATTATC TTCCGGGGTA TCTCGTGTGA AGAAGCTTAC
 129601 TATAGCCATC ATGGTCCTTT CATTGATAGC GATTTCTTTG GTAGCCTGTG
 129651 GCCTATTTTT AACGGGATCG GCACCTCTAC AGCTCTCGAT CTGGATTGCT
 129701 GCGAGTTGCA TTACCTTATC TATGTTAGTT TGTGCGTGTT GCGGTTATAA
 129751 GATTTCCAAT GCCTTAGAAA AAATAAGGT AGCGCATGAA AGCTGAGTTG
 129801 GACATTTTAT TGATTGAAAA ATCATGACTA CATTACCTAA GTACGTTCCC
 129851 CGTTCTCGAC AAAATCCCGA TACTCTGACC TTCCTAAAAC GGTATTCTAG
 129901 TGTCCTTCTC CATTGCGAGA ATTCTTTATC TTATCGGATT TTTGCGAAAG
 129951 TGCTTGCTAT TCTCCTCACT TCGTTAGCTG TAGCTTTCGC CGTGACTTTG
 130001 TTTTCTTGTT AAGGTTCTCA ACTGAGACTC TCGGCTCTCT ATATAGGTAT
 130051 AGCTCTTGCT ATTTGTGTTT TACTGACGAT CGTTGTTTAT TGTATCGCAA
 130101 GTAAAAATCGC CACAGCTTGC AAAAAGCCGC CTTCCATATC TCGAATTGAA
 130151 ATTGTTTAGA AGCATCTCTG TGTACAAAAG TTTACTAGAA ACTCGACTCT
 130201 AGGAAAGTTC CTAGAGAGAG CGACGCTGCG TTCGTGCTTT AGAAATACTT
 130251 GGCTGGTGTT TGGACGAAGA TTCTTTTAGT GGATTGGTCG TGTTTTCAAC
 130301 AACTTCAGTC TCTAGAGGAG CTCTTTGAAT CTTTGCTGAA GGTTTAGAAA
 130351 TATCAATACC TGTTAAGCTC ATAAAAGCCA TAGCAATGAG GCCTGTTGTA
 130401 ATGAAGGAGA TCCCATTCC CTGGAGGTTT TTGGGAATAT CAGAGTAGGC
 130451 GAGTTTTTCT TTGATAGTGG CTAAAATAAC AATAGCGAGC CACCACCCAC
 130501 ATCCCGCTCC TAAAGAGAAG ATCATCATAG GAATAAAAGG ATAACCTACGT
 130551 GTGATTCCGA AGAGCACACC CCCTAGGATC GCGCAGTTCA CAGCAATCAA
 130601 GGGAAGGAAG ATCCCTAAGG AGAGATATAG ATTCCTGGAG ACCTTTTCTA
 130651 AAAGAAGCTC TAAGATTTGC GTGAATGCCG CAATCACCAC GATGAAAATA
 130701 ATCAGCTCCA GAAAACCTAG GTTTACAGAA GCTAAAGATG GAGAGATCCA
 130751 AGTTAGAGCT TTAGGGCCCC TGATGAAAGC ATGGACAAAC CAGTTGATGC
 130801 TCCCTGTTAC AGTGAGAACA AGGGCTACGG ACATCCCCAA GCCATTGGCT
 130851 GTAGAAACCC TAGTAGAGCA AGCAAGGTAA CTACACATCC CCAAGAAATT

130901 CGCAAGAAGG ATATTCTGAA TAAAGGCTGC TTGTAGAAGA ATACCAAAGA
 130951 CATTAAGCCA AGTATACGCA CCTAACCACA TAAACTACCT TTTTCTCTTT
 131001 TTAGAGTCTC GAATGTTAAC AAGCCAAATC ATAATACCAA GTAGGAAAAA
 131051 AGCCGACGGT GCTAGCACCA TAAGACTTAA ATTTTGGTAT CCATCGGGGT
 131101 GGGTTTCGGA AGCATAAACA AATTGAGGGA TGATGCGAAA CCCATAAGA
 131151 GTTCCAAAAC CAAAGAGTTC TCTGATGACT CCAATGACAA GTAAGACCCA
 131201 GCCGTATCCT AAGCCAGAGG CAAACCCATC TAAGAACGCT GGAATAGGAG
 131251 TCACATGCCT AGCTAGACTT TCAGACCTTC CCATCACGAT GCAATTGGTG
 131301 ATGATAAGAC CCACAAAAAC AGAAAGTGTT TTGGAAATAT CAAAGAAAAA
 131351 AGCTTTTAAA AACTGGTCGA TAACAATCAC AAACAAGCTA ATGATAATTA
 131401 GCTGAGTAAT CATTCTCACA CTGTCAGGAG TGAACTTACG TAATAAGGAA
 131451 ACAAAGAAAG ACGAGCATCC TGTAACAATG CTGACAGCAA TTCCCATAGT
 131501 AATTGCCGTT TGTACTGTTG TTGTCACTGC CAGAGCCGAG CAAATCCCCA
 131551 AAATCGCAAT GAGAATTTGG TTGTTGCTCC ATAGAGGATC AAAGAAATAG
 131601 CTTTTATAGG ACTTTTTACT TGTCAATCGC CTGTTTTCTT TTCATGGGTT
 131651 AAATTAGAAA AATTTATAAG GAGCTGACGA TAGCAAGCCA GAGATTGTAC
 131701 ATAAGCTTCA GTGACACCGT TGCATGTTAA GGTGGCTCCA GAAATCCCAT
 131751 CAATAGCAGA AAGAGCTTTT GGAGAATCTC CCAAAGTAGT ACGCACGGAA
 131801 CCTTTAACTA CCTCAAGCCC TAGGTCTGTT GTTGCAAAAT TTGTAGTTCC
 131851 AGAAGAATCT TGTAGGAAGA TTTTCTTCCC ATAGAATTGC TCTTGCCATT
 131901 CGGGATTTGT AATATTTGCT CCTAAACCTG GAGTTTCTCC TTGTTGGTAC
 131951 CATGCGGTTT CCAATACAGT GTCACCGTCG TTTTTCACCTC CTAGATAGCC
 132001 ATGGATGGGG .CCCCAAAGGC CGAATCCTGA TATAGGGAAG ATCAAAGCTT
 132051 GAACTGTAGA AAGGTCTTTC GCAACGTCGG CTCCTGACAT ATTTTCTGTG
 132101 CGAGAGGTAT TCTCTAAAAT GACATAAAAG GGGAGGGGGG ATTGCTGACA
 132151 CGGAGGGCTT TCTTGATATT TCTCAAAAAA TTCAATGGGA TTCAGATTTT
 132201 TTTCTTCAAA AGAAAATACC TTGCCTTGGG CATCTGTAAG TAGAGGACGG

132251 ACAAAGCGCT CGGCATACAG CTCTAATTCA GGATAGGAAA CCTCAGAGAC
 132301 TTTTTTTGTA GCAACTTCAA GAAGTTGTGT TTTTTTATCG AAAGTCGCAG
 132351 GCACCCACTC TTTTTTTTCC TGAATTTGAA ATCTTCCTTT AAAATCTAAA
 132401 ATATGAGCAG CTAAAAGCAT TTGCTTATTG CGATCGAAAG TAGCAGCTTG
 132451 TTCCTGTATT GGGGAGAGCA CATAGTAGAT TGTGGATAAC AGCACTCCTG
 132501 CAAATAAGCT GAGGCCCAGG ATAAAGGAAA CGATGTACCA GGTTCGGTTT
 132551 ATGCGGACGG TATGTTTTGA AGAGCCTTTA GACATATTCT AGACTCCCCT
 132601 TTTTCTATAC TTTCTAACAG CAAAATAGTC GATAAGAGGG GCAAATACAT
 132651 TGCCCAGAAG GATCGCTAAC ATCACTCCCT CAGGATACGC AGGATGATA
 132701 AGACGAATCA CAATAGTCAT AAATCCTATA AAGAATCCGT AAATCCATTT
 132751 CCCTAATTTC ATAGTCGGCG ATGATACGGG ATCCGTAGCC ATAAAGACTA
 132801 AACCAAAAGC AAGTCCTCCG AGGAAAAGCT GCCGATAGGC GGAATGAAG
 132851 AATCGAGCAG GTGCCCAAGC TCCGTTTTGT CCCACGATGA GTACGCTGAT
 132901 AAACCTTAAAG AGCCAGCCTG TGAGAAAGGC TCCTATCCCA AAGGCTGCCA
 132951 TGGTTCCTCA AGAGGCAATG CCTGTAACAA TAAGGAATAT TGCACCCAAC
 133001 AGACAGGCGA AAGTGGAGGT CTCCCCCAGA GAACCTATAA TGTTCCTCCA
 133051 AAAGAGATTC CCAGCTGAGA ACTTCCCAAT CCCATAGATC ACATCGGTAA
 133101 TAGCATAGGC AGAATCGAAC TGTGTGGGAA GCAGCCCCAA TCCTCCCTCA
 133151 GCAACAGGAG CTGTAACAAA CGTTTGAAGT TGTGTAAGAG TGAGATTATC
 133201 TAAAACCCAA CCAGGATGCG TCTCTGTCCA AAGAGAAAAT TGTGAGTGAA
 133251 TGACATCTTG AGTAGGGACG TGAGGAATGT GAAGCATATT TGCAGCAATC
 133301 GCATCGACAT GCAGACGCTT TACAGAGGGA GGTGTCGAAT TTAGAGTTTG
 133351 TAGGCAGGTA GACTGTGAAA ATCCATCAAT GAGTACTTTT CCTGTCGAGG
 133401 AGTTCATCTT CATGAGGCTA TCTTTAATCA CTCCGGGGTT GCTTCCTACC
 133451 CAAACGTCAC CACTCATCTT TGCTGGAAAC GTAAAAAATA AGAATGCCCT
 133501 TCCTGATAGA GCAGGATTGA GGATGTTTCA CCCTGTGCCT CCGAAGAGCT
 133551 CTTTACTGAC AACAATACCA AAGGCGATCC CTAAGGCTGC CATCCAGTAA

133601 GGAATTGTCG GAGGGAGAGT AAGGGGATAG AGGATTCCGG TTAGTAGCAG
133651 TCCTTCTGCG ATTTTATGCC CACGAACTAC AGCAAATAGG ACCTCACAAG
133701 TACCCCCGAC AACATAGCTA ATCGTAAGTA GAGGAATAAA GATCTTAAGT
133751 CCTTCCCAAA GGATAGGAAC TATATGGATC TCTTTGTAAA CAAAGGATAA
133801 ATAACTACCA AATCCAGAAA TATGTAAGAA TTGCTCCATC AGCACAGGAT
133851 TGCCTGAGCT ATAAACGATA GATTGAAGTC CTGAATTCCA GATCGCAACA
133901 AAGGTCGCGG GAAACAAAGC GATAACAACA AGCATCATCC AACGCTTAAC
133951 ATCTACAGAA TCGCGGATGA AAGGAGGCTT GGAAGGGGTT TCAATAGGTT
134001 CGTAACAAAA TGTATCTATC GCATCGACAA TGGGAGTAAA GCGCTGATAC
134051 TTGTCTTGTT GACATAGTTT CCAAAGAGAA TTTATGAATT TTTTGAGCAT
134101 TGTGATTGAG AGAAAAGTTG AAGCTTCTAC ACGTTCAAAA ACGTAGCAAA
134151 CTGCTTAAAA TTTTAGGAAT AAAAATTTTC ATGATTCAAA AATAGATGGT
134201 ACTTTTTTCG TTGCTGTTTC CAAAGTTATG TTATGGCTGT CAAGCTCCAG
134251 GAGCCTACTT TTGTTCCAAC TGCTTGGAAG AACCTCTCGT AGAAGATAGA
134301 GAAGGGCGTT GTCTACATTG TTTTCGTTAT CTTGGTTCTT CCGAAACACG
134351 TCTATGTAGC CAGTGTTCAC CCTCTTCACA ACTTCAAGCT TTCAGCTTGT
134401 ACCTTCCTTC GCAAACGGCC CTCTCGGTAT ATGCTCGTGC TTGTGAAGGT
134451 AAGCGACCCG CTCTGCAGTT TTTTCTAAG AGTATCGCCT TTGAGCTAGC
134501 TTCACTGGAT GAGACTCCGA GTTGTATTGC CTATATAACA TCGACAATTT
134551 CTAGGAAAAT CGTAGTAGAA GTTGCTAAAC TAGAAAAGCT TTTACGCATT
134601 CCCTTGTTGGC CGTGGCTTCC TAAGAAAAGA CAAATAGAAA AACTTCCTAA
134651 AGGGGAAGGT ATCTGCTTTT TGTCGGCCTA TCCTTTATCA CAAAAATGGA
134701 TGCAAACATAT CGTTGGAGGG AGTGCATCAC CTCTAGTATC TATAAGTCTC
134751 TTTCTCTCTC AGAATGATCA GTAATTCCTG CAATTGCAAG GTAACCAAGA
134801 ATACGTACGC CCTCATCAAC ACACAACGTG TTCGCTTGGA TGTCTCCTTT
134851 AATGATTGCG CCTCCACGGA GTTCGACTTT TCCAGATACT GTGATATTTT
134901 CTTCTACAAC CCCTTCAATA ATGGCTTCTT GTAGCTGAAT ATCTGCCTTT

134951 ACCACTCCTT TAGGACCGAT AATAATTTTT CCTTTTGAGA CTAAAATGCC
135001 TTCAAAAGTT CCGTCAATAC GTAGGAGACG TTCAAAAGCA AGTTCTCCTT
135051 TAAAGGTGAC GCCTTCTCCT AAGGTAGTTT CAGGTCTTTC AAGAGGGAGT
135101 AGAGATTCTG TTCTTGGAGT TGAGGACCAT TGAGGAAGAG AAGATTCTTC
135151 AGTTAAATTG TGATTCAAAG GGCGAGCTTC CGAAGCTTTA GGGTTGTCAA
135201 AAAGACTTGG AGGGGTCTCT GGGCGCTCGG ATCTTGAATA TGGCGAGTAG
135251 CTGGAAGGTG AAGAAGTTTC TTCTTCGTAA AGTGTTTGCA CATCTTCAAA
135301 AGGACCTTTT CCTGTTCTAC GGAACATGGG ACACCCCTA AATTAACCAA
135351 CAATATGATT TTTACAGAGA TTTACTTGAC GCTTAAGAGT TTCGTCATTT
135401 TGTAATTTAT GTTCTATAGT TTTACAGGCA TAAAGTACTG TCGAATGAGT
135451 TTTACCAAAA GCAGCTCCTA TTGCAACTAA AGAATCTGTA ATAAGAGTTT
135501 TTGCTAAATA CATAGCAATT TGCCGAGCTA ACACAAGATC TTTAGAGCGT
135551 GAGTTTCCCT TAAGATCATT CAGCTTTACT TGAATACTG TAGCAACACT
135601 TTTTAAGATC GTTCTACAG AAATTTTTTG TTTTGTTGGA GAACGGAAGA
135651 GCTCTTTTAG AGTTTCTCGG ACTGTAGTTT CTGTAAGAGA CTTGCCGAAA
135701 AGACGACAAT AGGCAGTCAG CTTGTTGATA GCTCCTTCCA ATTGACGGAC
135751 ATTGCCATAG ATGTGATCCG CAATATAAAA TGCCATTTCA TTAGGAATGA
135801 GCAATCCTTT TTGCTCCGCC TTGTGCTGTA AAATCGCAAC CCGAGTTTCT
135851 AAATCAGGGA TGCCGACGTG AGCAACCACT CCCATTCCA TTCTAGCAAT
135901 GATACGCTCG GAAAGTTTGA GCTGACTTGG AGGTTTATCA CTGGTAATTA
135951 CAATTGCTT ACTCAGGTTG ATCAAAGTCT CAAAGGTATT GCAAAACTCT
136001 TCTTCAAAAT TTTGGCGATT CTGTAAAAAT TGAATATCAT CAACAAGAAG
136051 TAAATCTAGG GAACGATAAA AATTTTTCAT TTTATCAACA GACTTGATT
136101 TGAGATGGTA GACAAGATCG TTGATAAACG CTTCTGTAGT GATGCAATGG
136151 ATGCGTAGAT TTTTATGATG TTCTCTTACG TAGTGACCTA CGGCATGAAG
136201 TAAATGCGTT TTGCCTAATC CCACACCCCC ATGGATGAAT AAAGGGTTGT
136251 AGGAGCGGCC AGGTTTCCCA GCAATACCTA CAGCTGCAGA CTTCACAAAT

136301 TGATTTGAGG GACCTTCAAT GAAATTATCA AAGCGATAGG AGAGATTCAG
 136351 CTTTAATTCA AAATCTTTAG TTTCTTCAA GACCTCAGAA ATTCCTTCGT
 136401 TTGATTCTTT TTGAGAAGCC ACGGGGGCTG AAGGTTTCTT GTGTTCTGCA
 136451 ACTACAAATT CTAAAGCAGG CTCTCCATGA ACATCTAAGG GGACAAAAGA
 136501 ACAGAGGTCT CTTTTGTAGT TATCAAGAAG ATAATTTTGT AAAAAATGT
 136551 TGGGGACTTC TAAGCGAATT TTCTCTTGAG TTTCTTCAAG AACTTGAATA
 136601 GGAGAAATCC AATTTTCAA AGCCGTTTTC GAGCAACGTG TCTTAACATA
 136651 ATTTAAAAAC TGTTCCCAAG TAGTGCACTC GTTACAGGTT AACATGCCGC
 136701 TCTCTTTATT TATAAAGCTT TCCCAAATAC AATCGACCCA TCCCATGAGT
 136751 GATGCGGAGA AAATCTCATT GCATCTTGAC TATTCATTGC ACTGAGTGCA
 136801 ATGACTCTTC GCGTTCAGAT TGGATCCTGC ATTCCCTGGG TGAAGTTCCT
 136851 GTTAGGAATG GATCCTATAC ACCCTTCCTC CGCAGGTAAA CGCGGTACGC
 136901 TCTGCCGATA AAATAACTCT ACCGATTACT AGGTTTAAAG GCAAATTGGA
 136951 TCGTTGGTTT CGTTAGGCAA TAAGGAACCA CAAATTCAGG AAAAAATAAT
 137001 TATGAAATTT TGTAATAAAA ATGGAAAAAG AACTAAAGAA ATCCGGAGTT
 137051 CTTCAATCAC GAAATACGTT TTCTATAGGA GAAAAAATTA ACGAACTAAC
 137101 GCAGCATTTT TTTTGGTTGC TTTACTATAA CTCATCAATA GAGCTTCAGC
 137151 ATCATTAGCA ATTGCTGGTA TGGGACAGGA TGAAAGGTAG GTGGCAATGG
 137201 CAGTAGCTTC TTCAATTCGT CCCAAACAGA AGAGAGCTTT TGTTTTATTT
 137251 AAGAGTGTAG GCAGATGATC TCCTTGCATG CGGAGTGCCT GATCTAAAAC
 137301 AGCAAGCGCC TGACTATTTT CACCAATTTG GAGATAAAGA CCTCCAAGAG
 137351 TTTGATGATC ATAGATACTT AAAGGATCTA AGATCACTAG AGCTTCAAAA
 137401 AAAAGAATCG CTTTTGAATA ATGCCCTTGG CGTAGAAAAG AATATCCTGA
 137451 GATTCTGAGT TCTTCTAACT CATCATCTCC CCAGCCTAAG ATTGCTTTCC
 137501 ATTCATTATC CAACATACAT TATCCTTGAA CAAATTGAAA GATACGAGAG
 137551 ATCACATAGT CTATCATCAC ATTTGTAGAC TTGACTCCAA GATCTAGAGC
 137601 TTTTAAGAGC ACTTTCCTT GTTGTACTCT AGGATCGTCA TCGTCTTCTT

137651 CAGGATCTTC GTCTTCCTCT TCTTCTTTAT CTTTGTAAGA CAAAAAGGGA
 137701 GTAATTTGAC TGCGATAGGA AAACCTCCCT CGAGTGAGAA CTTTAAAAA
 137751 TGAGGAGATT TTTTCTATGT CTTTCTCTTG TTGGTCTGGA GATCCTAAAG
 137801 AAGGTGCCAG GTAGGGTGTG GAAAAACGCT GTTTGTAAAA ATTATTTGGA
 137851 GGGGAAAAAC ATGCCCAGTG CGACTTTTGA TTTGTCTGTA AAAGAGACGT
 137901 CAAAGCCGAA GGCTTGCGGT TCATATCCAA AATTGCGCA TGCTTGCGAA
 137951 CATCACGAAT GGAGATGCCT TCCATCTGGA TTTCTTTGCG AAAGTCGCTG
 138001 ACTATCCTAT TATTGGAAGC ATGTTGCTCA TATATAGACG TGCTATAATT
 138051 AAAAATTTCT ACCATGGCGA GGCCTTAAAA AACCGTCTCT ATCTCTAGAT
 138101 GATAGTGCAG TCTAGGAAAA AGTCAATAGT CTTGATCAGA AGCCTGAACC
 138151 TTTTCTCCTT ATCGAAGGCA CTTAAGAAAA AAGCTCACCC CTATCAAAAA
 138201 ATTTAGAGTG GCGAACTAGC GAGAATTTAA GATAAGGGAA GGCTTTCTTC
 138251 TTTCTATGGA ATCCTGTATC TTTGTCCTTA TTTGAAGCCA AGAGCTTAGG
 138301 AAATTCTTAT GTCCGAACGT GCGCATATTC CCGTATTAGT TGAAGAATGT
 138351 TTAGCTTTAT TTGCTCAACG TCCTCCACAG ACTTTTCGAG ATGTCACCTT
 138401 AGGAGCTGGA GGACATGCGT ATGCTTTTCT TGAGGCGTAT CCCTCTCTAA
 138451 CTTGTTATGA TGGCTCCGAT CGAGATCTTC AGGCTTTGGC AATTGCAGAA
 138501 AAACGTTTGG AGACCTTTCA AGATAGAGTC TCCTTTTCCC ACGCCTCTTT
 138551 TGAAGATCTT GCGAACCAAC CCACTCCACG TCTTTATGAC GGAGTTCTTG
 138601 CAGATTTAGG AGTCTCTTCT ATGCAGCTGG ATACTCTATC CCGAGGGTTT
 138651 AGCTTTCAAG GGGAAAAAGA AGAGTTGGAT ATGCGTATGG ATCAAACGCA
 138701 AGAGCTTTCC GCTAGCGATG TCCTGAACTC CCTAAAAGAA GAAGAACTAG
 138751 GGAGAATTTT TCGTGAATAT GGAGAGGAAC CACAATGGAA ATCTGCAGCT
 138801 AAAGCTGTTG TCCATTTTCG TAAGCATAAA AAAATTCTTT CGATCCAGGA
 138851 TGTAAGAGAA GCTCTTCTTG GCGTTTTCCC TCACTATCGT TTTTATAGAA
 138901 AAATACATCC ACTCACCTTG ATTTTTCAGG CTCTACGTGT TTATGTGAAT
 138951 GGAGAGGATA GACAATTGAA AAGTTTACTA ACATCTGCTA TATCTTGGCT

139001 GGCTCCTCAG GGACGGCTTG TCATTATTTT TTTTGTAGC TCTGAGGATC
139051 GTCCTGTGAA GTGGTTTTTT AAAGAGGCGG AAGCTTCTGG CCTGGGGAAG
139101 GTAATCACAA AGAAAGTGAT CCAACCTACC TACCAAGAAG TACGAAGAAA
139151 TCCTAGATCG AGATCAGCAA AACTACGGTG TTTTGAAAAA GCTTCCCAAT
139201 GAACAAAAGT CGTTTTTTTAC GTTTATGCTG CTGTCTATGC TTTTGTGGAA
139251 GTCTCTTTTA TTTCTATATT AATAAGCAGA ACTCGCTGAC GAAATTACGC
139301 CTCGAAATTC CTTGTTTATC TGTACGCTTG CGTCAGCTTG AGCAGCAAAA
139351 TATTTCTTTA CGTTTTTTTAA TTGATAAAAT AGAAAGACCT GATCATTTGA
139401 TGGAAATAGC AGCTCTTCCC GAATACCAAT ATTTGGAATA TCCCTCAGAA
139451 GAAAGTATCA GTCTTTTATC CTATGAGCTA CCGTAAACGT TCGACTCTAA
139501 TTGTTCTAGG AGTGTTTGCT CTTTATGCTC TTCTAGTATT GCGTTATTAT
139551 AAAATTCAAA TTTGTGAAGG AGACCACTGG GCCGCAGAAG CTCTCGGGCA
139601 ACACGAATTT TGTGTCCGTG ATCCTTTTCG AAGGGGCACC TTTTGTGCTA
139651 ACACGACAGT ACGTAAGGGA GACAAAGACC TTCAGCAGCC TTTCGCTGTC
139701 GATATTACAA AATTTACCTT TTGTGCAGAT CCTTTAGCTA TTCCCGAATG
139751 TCATCGTGAT GAGATCATCC AAGGGATTCT CCAATTTATT GAGGGGCAGA
139801 CCTACGACGA CCTCTCCCTA AAGTTAGATA AGAAATCTCG GTATTGTAAG
139851 CTGTATCCTT TATTAGATGT TTCTGTCCAT GACCGGCTAT CCCTTTGGTG
139901 GAAAGGATAT GCAACAAAGC ATCGCTTACC AACAAACGCC CTATTTTTTA
139951 TTACGGACTA CCAACGCTCG TATCCTTTTG GGAAGCTCCT TGGACAAGTT
140001 CTCCATACCT TAAGAGAAAT TAAGGATGAG AAAACAGGAA AAGCCTTTCC
140051 CACAGGCGGG ATGGAGGCGT ACTTTAATCA TATTCTGGAA GGGGACGTTG
140101 GAGAGAGAAA GCTGTTGCGT TCTCCTTTGA ACCGTTTAGA TACGAATCGT
140151 GTTATCAAAC TGCCTAAAGA TGGCTCTGAT ATCTACCTTA CGATCAATCC
140201 TGTGATCCAG ACCATTGCAG AGGAAGAAGT CGAACGGGGC GTGCTAGAAG
140251 CTAAAGCCCA GGGGGGTAGG CTCATTCTAA TGAAGTCCCA AACAGGAGAG
140301 ATTCTTGAC TGGCTCAATA TCCGTTTTTC GATCCCACAA ATTATAAGGA

140351 ATACTTCAAT AACAAAGAGC GCATCGAACA TACGAAGGTA TCTTTTGTGA
140401 GCGATGTTTT TGAACCCGGG TCGATCATGA AACCTTTGAC TGTGGCGATT
140451 GCTTTACAAG CTAACGAAGA GGCTAGCTTA AAATCGCAGA AAAAGATTTT
140501 TGATCCTGAA GAACCTATCG ATGTGACCAG GACACTCTTC CCTGGACGAA
140551 AAGGATCTCC GCTTAAGGAT ATTTCTAGAA ACTCTCAATT GAATATGTAC
140601 ATGGCTATCC AGAAATCTTC GAATGTCTAT GTAGCTCAGC TGGCTGACCG
140651 CATCATACAA TCTTTAGGAG TGGCCTGGTA CCAACAGAAG TTGCTAGCTC
140701 TGGGATTTGG AAGAAAAACA GGGATCGAGC TTCCAGTGA GGCCTCTGGT
140751 TTGGTGCCTT CTCCCCATCG TTTCCATATT AATGGTCCC TGGAAATGGTC
140801 CTTATCTACT CCATATTCTT TGGCTATGGG ATATAATATT TTGGCAACAG
140851 GGATACAAAT GGTTCAGCC TACGCTATCC TTGCAAACGG AGGTTATGCC
140901 GTCCGCCCCA CTTTAGTAAA AAAGATCGTC TCTGCTTCAG GAGAGGAATA
140951 TCATCTTCCT ACTAAAGAGA AGACACGACT CTTTTCAGAA GAAATTACTA
141001 GAGAAGTTGT TCGTGCCATG CGTTTTACAA CGTTACCCGG AGGTTCCGGA
141051 TTTCGAGCCT CTCCTAAGCA TCACTCTAGT GCTGGGAAA CAGGAACTAC
141101 AGAAAAGATG ATTCATGGAA AATATGATAA ACGCCGTCAT ATTGCTTCTT
141151 TTATAGGTTT TACTCCCGTA GAGAGCTCGG AGGGAAATTT CCCACCTTTA
141201 GTGATGCTCG TCTCCATAGA TGATCCTGAA TATGGTTTGC GAGCCGACGG
141251 CACGAAAAAT TATATGGGGG GGCCTTGTGC GGCACCCATT TTTTCTAGGG
141301 TTGCTGACCG CACACTCCTC TATTTAGGGA TTCTTCCAGA CAAGAAGCTA
141351 AGAAATGCG ACGAAGAAGC TGCTGCATTA AAGCGTCTCT ATGAAGAATG
141401 GAATCGTTCT CCGAAACAAG GGGGAACGAG GTGAGGATCT CTATTTCCAT
141451 CTTGCTATAG ACTTTTACCG TTGAGCAAAG ACTCTCTATC AGAGAGCCCG
141501 TCTCCTCTTT ATCCTCTATG AGTAGTTTAT GTTATGGCTA GGGTAGGTCC
141551 TAAACTATAG AAATAACTTT AGCTTTCTTC CCCTAAATAA GAGACCAAAG
141601 TCTTGATGAG ACGGTCTATT GAAGTTTATG GAAGGGGGAG GTAAGGCTGT
141651 GTGTTTGGGG ATTTAGATTT GGGATAAAGG AGGCTTCTGT TCGTAGAAAC

141701 AGGAGAGCGA AATTTTATAT TTCAGAGAAG AGTAAGAACT TTATGGACAG
 141751 TTTTTTGTGA TTGCTTGAT ACTATCTTGA TTGAATTTTT TGTCGACCTA
 141801 CGAGTAAAGA AATCCTTTAA GCATTTTTTA AAAATCAGAG TGAGAGCATG
 141851 CCCCTAGAGG GCTTTTTATG AAAAAAGTTG TTTTCAATA GTCCCTGGAG
 141901 CGTAAATGGA TTTAAAAGAG TTACTCCATG GGGTTCAAGC TAAAATCTAT
 141951 GGGAAAGTTC GCCCTCTTGA AGTGCGCAAC TTGACACGTG ATTCCCGTTG
 142001 TGTGAGTGTT GGCGACATTT TTATAGCCCA TAAGGGACAG CGCTACGACG
 142051 GAAATGATTT TGCTGTCGAT GCTTTAGCTA ATGGAGCAAT TGCCATTGCT
 142101 TCTTCACTAT ACAATCCGTT TCTTCCGTT GTTCAGATCA TCACTCCTAA
 142151 TCTCGAAGAA TTAGAGGCTG AGCTTCTGC AAAGTATTAC GAATACCCTT
 142201 CAAGTAAGCT CCATACCATT GGGGTGACTG GAACCAATGG GAAACTACA
 142251 GTTACATGTT TGATTAAAGC TTTATTGGAT AGCTATCAAA AACCTTCAGG
 142301 GCTTTTAGGA ACCATAGAGC ATATCTTAGG AGAGGGGGTG ATTAAAGATG
 142351 GGTTTACTAC ACCTACACCC GCTCTTTTAC AGAAGTATTT AGCCACTATG
 142401 GTACGTCAAA ATAGAGACGC TGTGTTATG GAAGTCTCTT CTATAGGACT
 142451 TGCCTCTGGA AGAGTAGCCT ATACCAATTT TGATACAGCA GTTCTGACTA
 142501 ATATTACCTT AGATCATCTC GATTTTCATG GCACATTTGA AACCTATGTT
 142551 GCGGCGAAAG CCAAGCTTTT CTCTCTCGTG CCCCTTCGG GAATGGTTGT
 142601 TATCAACACA GACTCTCCCT ACGCTTCTCA GTGTATTGAG AGTGCAAAGG
 142651 CACCGGTCAT CACTTATGGT ATAGAGAGTG CTGCTGACTA CCGAGCCACC
 142701 GATATCCAAC TTTCTTCTC GGAACAAAG TATACCTTGG TGTACGGGGA
 142751 CCAAAAAATT GCGTGCTCTT CCTCATTTAT TGGAAAGTAC AACGTCTATA
 142801 ACCTACTTGC TGCGATCTCT ACAGTACATG CAAGTTTGCG TTGCGATCTT
 142851 GAAGATTGTC TAGAAAAGAT AGGCTTGTGT CAACCTCCTC CAGGTCGTTT
 142901 GGATCCTGTA CTTATGGGTC CCTGCCCTGT ATATATTGAT TATGCACACA
 142951 CCCCCGATGC TTTAGACAAT GTCTTAACAG GATTGCATGA GTTACTTCCT
 143001 GAGGGGGGAA GACTGATTGT TGTTTTTGGT TGCGGTGGAG ATAGAGATCG

143051 CAGTAAACGG AAGTTGATGG CCCAGGTGGT AGAGCGTTAT GGTTTTGCTG
 143101 TTGTAAGTTC AGATAACCTT AGGAGCGAGC CTCCTGAAGA TATTGTGAAT
 143151 GAAATTTGTG ATGGGTTTTT TTCAAAAAAC TATTTTCATCG AAATCGACAG
 143201 AAAACAAGCA ATTACATATG CTCTGTCTAT TGCCTCAGAT AGAGATATAG
 143251 TGTTAATAGC GGGAAAAGGG CATGAAGCTT ACCAAATATT TAAACACCAA
 143301 ACAGTTGCGT TCGATGATAA GCAGACTGTT TGTGAGGTAC TCGCTTCCTA
 143351 TGTCTAAGCA ACTGTCGTTT TTTGCTTTAT GTGTGTTAGG AAGTCACCCG
 143401 ATTTTTGCTC AAACACCGAA TCCTCCTCAG CGTGTACGAC GCAGTGAGGT
 143451 TATATTTATA GATCCTGGAC ACGGGGAAA AGATCAAGGC ACGGCAAGTA
 143501 AGGAAGTTC TATGAAGAG AAGTCCCTGA CCCTGTCTCT TGCTTTGACG
 143551 GTTCAAAGTT ACTTAAAGCG GATGGGTTAT AAACCTCAGC TAACCCGATC
 143601 TTCTGATGTA TACGTTGACT TAGGGAAACG CGTTGCTTTG TCGAACCGTG
 143651 GGCAGGGGGA TGTCTTTATC AGCATCCACT GTAATCATTC TTCAAACGCA
 143701 GCAGCCTTTG GCACCGAAGT ATATTTTTAT AATGGTAAGG TCGGATCTCC
 143751 GACTAGGAAT CGCATGTCAG AAGTACTGGG AAAAAACATT TTAGCTGCTA
 143801 TGGAAAAAAA TGGCATTTTG AAGTCTCGAG GTTTGAAAAC TGCGAACTTT
 143851 GTTGTGATTA GAGATACTTC TATGCCTGCA GTTTTGGTGG AAACCGGGTT
 143901 TTTATCCAAT AGTCGTGAAC GTGCGGCCCT GCAAGATGCT CGCTATCGTA
 143951 TGCATGTAGC GAAAGGCATC GCCGAGGGAG TTCATAATTT TCTTTCTGGA
 144001 CCTAGTTTTC AGAAACCAAA ACAGAATATC GCTAAAATAC GTAAACCACA
 144051 GATACAAGCA AATTAGTACT TTAGGAGTTA AAGGCAAAAA ATCGTCCTCG
 144101 ATGCGATTCG AACGCATGGC CTGCTGCTTA GGAGGCAACC GCTCTATCCT
 144151 GCTGAGCTAC GAGGACGCAA AGACCAGCAC TTTACCAAGT TAGATTAAAG
 144201 AAGTCACGTG TTAGATGCAC GCTAGCAATT TAGGGGAAGT TTTTCTCAAG
 144251 ATGTGGGAAT GATTTTTCTA GGTTCAGAA ATATAGTTAT TCGCATTAAT
 144301 CGATATGGTT TATAGTGATT GCGCATTTTT TAAAAATGT CTTGAATCCA
 144351 AAGGATGAAT AGATATGATG AGCTTAGACT TCAAAAGTTT ATTATTTACG

144401 GTTACTCAAC GAGCTACTTT AGGGCACTTT AATAGGAGGC ATTGTCTAAT
144451 ATGGCTACCA TGACAAAGAA GAAACTAATC AGCACGATCT CACAAGATCA
144501 CAAAATTCAT CCTAATCACG TACGTACCGT GATTCAGAAT TTTCTAGATA
144551 AAATGACCGA CGCCTTGGTT AAAGGTGACA GGCTTGAGTT TAGAGATTTT
144601 GGTGTGTTGC AAGTAGTAGA AAGAAAACCA AAGGTAGGAC GTAATCCTAA
144651 GAATGCAGCA GTCCCCATTC ATATTCCTGC TAGACGCGCT GTAAAGTTTA
144701 CTCCAGGGAA AAGAATGAAG CGCTTGATAG AAACCTCCGAA TAAGCATTCT
144751 TAATTCTTGT AGTCTTCTTT GTCTCAGTTG TTAGAGTCAG ACCGGTTTTT
144801 TACCGGGCTT GACTCTAATT TTTGTTATTA TTATCGTTTG GTGCAATGCT
144851 TTTCTGATCA AATTGTGCGT GATAATGGGG CTGCAATCCA GGTTACAACA
144901 TTGTATAGAA GTGTCCCAGA ATTCGAACTT TGATTACAA GTAAAACAGT
144951 TTATCTATGC GTGCCAAGAT AAGACATTAA GGCAGTCTGT ACTCAAGATT
145001 TTCCGCTACC ATCCTTTACT AAAAATTCAT GATATTGCTC GGGCCGTCTA
145051 TCTTTTGATG GCCTTAGAAG AAGGCGAGGA TTTAGGCTTA AGCTTTTTAA
145101 ATGTACAGCA GTACCCTTCA GGTGCTGTAG AACTGTTTTT TTGTGGGGGA
145151 TTTCTTGGGA AAGGATTACC TTATCCTGCA GAACATGCGG AATTTGGCCT
145201 ACTCCTGTGA CAGATCGCAG AGTTTTATGA AGAGAGTCAG GCATACGTCT
145251 CTAAAATGAG TCATTTTCAA CAGGCACTCT TTGATCACCA AGGGAGCGTC
145301 TTTCCCTCTC TCTGGAGCCA GGAGAACTCT CGACTCCTAA AAGAAAAGAC
145351 AACTCTTAGC CAATCGTTTC TCTTCCAATT AGGAATGCAA ATTCACCCAG
145401 AATACAGTCT TGAGGATCCT GCACTAGGGT TCTGGATGCA AAGAACGCGT
145451 TCTTCATCCG CTTTGTAGC CGCTTCAGGA TGTCAAAGTA GCTTGGGAGC
145501 GTATTCCTCA GGGGATGTGC GTGTTATCGC TTATGGACCT TGCTCTGGAG
145551 ACATTAGTGA TTGTTATTAT TTTGGATGTT GTGGAATCGC TAAAGAGTTC
145601 GTGTGCCAAA AATCTCACCA AACTACAGAG ATTTCTTTTC TCACCTCTAC
145651 AGGAAAGCCT CATCCCAGAA ATACGGGATT TTCCTACCTT CGAGATTCCCT
145701 ATGTACATCT GCCGATCCGC TGTAAGATCA CTATTTCCGA CAAGCAATAT

145751 CGCGTGCACG CTGCGTTGGC TGAGGCCACC TCTGCCATGA CGTTTTCTAT
 145801 TTTCTGTAAG GGGAAGAATT GTCAGGTTGT TGACGGCCCT CGCTTGCCT
 145851 CCTGTTCCCT AGATTCTTAT AAAGGTCCCG GAAACGACAT TATGATTCTT
 145901 GGGGAAAATG ACGCAATCAA CATTGTTTCT GCAAGTCCCT ATATGGAAAT
 145951 TTTTGCTTTG CAAGGCAAAG AAAAATTTTG GAATGCAGAC TTTTGTATTA
 146001 ATATTCCTTA CAAAGAAGAG GCGTCATGT TAATTTTGA AAAAAAGTG
 146051 ACCTCTGAGA AAGGAAGATT CTTTACGAAG ATGAATTAAT TTTGGGTCTG
 146101 TAATTGTGTT TAAGAATTGT TTGTATTAA ATGATTCTTT TTATACGAGG
 146151 AGAGCACATT CTAATGGAAC TTCTTCCACA CGAAAAACAA GTAGTTGAAT
 146201 ATGAAAAGGC TATAGCCGAA TTAAAGAAA AAAATAAGAA AAATTCTCTC
 146251 TTATCTTCTT CAGAGATTCA GAAATTGGAA AAGCGTTAG ATAAATTAAA
 146301 AGAAAAGATC TATTCGGATT TGAATCCTTG GGAGCGTGTA CAAATATGTC
 146351 GCCACCCTTC GCGTCCCCGT ACTGTCAACT ATATTGAAGG GATGTGTGAG
 146401 GAGTTTGTG AGCTTTGTGG AGATCGCACC TTCCGAGATG ATCCCGCAGT
 146451 TGTGTTGGC TTTGTAAAAA TCCAGGGTCA GCGTTTTGTC CTTATTGGCC
 146501 AAGAAAAGGG ATGCGATACA GCGTCACGCC TTCATAGGAA CTTCCGTATG
 146551 TTATGTCCCG AGGGTTTCAG AAAAGCCCTT CGCTTAGGAA AACTCGCTGA
 146601 AAAGTTTGGC TTGCCTGTGG TCTTTCTGTG CGATACCCCA GGAGCATATC
 146651 CTGGATTGAC TGCTGAAGAG AGAGGACAAG GATGGGCAAT TGCCAAAAAT
 146701 CTTTTTGAGC TCTCAAGACT TGCCACTCCC GTGATTATTG TCGTTATCGG
 146751 TGAGGGATGT TCAGGTGGAG CTTTGGGCAT GGCTGTAGGT GATTCTGTAG
 146801 CTATGTTAGA GCATTCCCTAT TATTCTGTAA TTTCCCCAGA AGGATGCGCC
 146851 TCCATTCTTT GGAAAGATCC TAAGAAAAAT AGCGAAGCAG CTTCCATGTT
 146901 GAAAATGCAT GGAGAAAAC TAAACAATT TGGCATTATC GATACTGTTA
 146951 TCAAAGAGCC CATTTGGGGA GCTCACCACG ATCCTGCATT GGTATATAGC
 147001 AATGTTGAG AGTTTATCAT CCAAGAGTGG TTACGATTAA AAGATCTAGC
 147051 TATAGAAGAG CTGTTGGAGA AACGGTACGA AAAATTTGCG TCTATAGGTC

147101 TTTATGAAAC TACTTCTGAA AGCGGTCCTG AGGCATAAAA ATCATCTCGT
147151 TATATTAGGC TGTTCTCTAC TCGCAATTTT AGGACTTACC TTTTCATCTC
147201 AGATGGAGAT TTTTCTTTA GGGATGATTG CTAAACAGG CCCCAGCGCC
147251 TTTTACTTT TTGGACGTAA GGAATCTGGA AACTTGTAAGGTTTCAGA
147301 ACTAAGTCAG AAAGATATTT TAGAGAATTG GCAGGCAATT AGTAAGGATT
147351 CAGAGACACT TACAGTCTCT GATGCCACGA CATACATCGC CGAACATGGG
147401 AAAAGCACAG CCTCTCTGAC GAGCAAGCTC TCTAAGTTTG TCCGTAAC TA
147451 CATCGATGTG AGCCGCTTTC GAGGACTGGC AATCTTCTTA ATCTGCGTTG
147501 CTATTTTAA AGCAGTCACC TTATTTTCC AACGTTTCCT TGGGCAAGTC
147551 GTTGCTATAC GGGTAAGCCG AGACTTACGT CAGGACTACT TTAAGGCCCT
147601 ACAACAAC TC CCCATGACCT TCTTCCATGA TCATGATATC GGTAATTTAA
147651 GTAATCGTGT CATGACAGAT TCTGCAAGCA TTGCCTTAGC AGTAAACTCT
147701 TTAATGATTA ACTACATTCA AGCCCCAATT ACCTTCATAT TGACATTGGG
147751 AGTCTGTCTG TCGATTTCAT GGAAGTTTTC AATTCTTATT TGTGTTGCCT
147801 TTCCTATCTT TATCCTTCCC ATTGTCGTGA TCGCTAGAAA GATCAAAAAT
147851 TTAGCAAAAC GTATTCAAAA GAGTCAGGAT TCATTTTCCT CCGTCTTTTA
147901 TGATTTTCTT GCTGGGGTTA TGACAGTAAA AGTCTTTCGT ACAGAAAAAT
147951 TTGCCTTCAC AAAATATTGT GAGCATAACA ATAAGATTTC TGCTTTAGAG
148001 GAGAAAAGTG CTGCTTACGG TTTGCTTCCA CGACCCCTCC TGCATACCAT
148051 AGCTTCTTTA TTTTGTGCTT TTGTCGTCGT TATCGGAATT TATAAATTG
148101 CTATTCCTCC CGAAGAACTT ATCGTATTTT GTGGTTTGCT CTACCTAATC
148151 TACGACCCTA TTAAGAAGTT CGGGGATGAA AATACCTCCA TCATGAGGGG
148201 ATGTGCTGCT GCGGAGAGAT TTTATGAAGT CTTGAATCAC CCCGATCTTC
148251 ATAGTCAAAA AGAAAGAGAA ATCGAGTTCC TTGGACTTTC TAATACAATC
148301 ACATTCGAGA ATGTTTCCTT CGGCTATCAG GAAGATAAGC ACATCCTCAA
148351 AAATCTAAGC TTTACCTTAC ATAAAGGCGA AGCTCTAGGC ATTGTAGGAC
148401 CTACAGGATC TGGAAAAACA ACACTTGTTA AATTACTTCC TAGGCTCTAC

148451 GAAGTCTCCC AAGGAAAGAT TCTTATCGAC TCTCTTCCTA TTACGGAATA
148501 TAACAAAGGG TCCTTAAGGA ATCACATCGC CTGTGTATTA CAGAATCCTT
148551 TCTTATTCTA TGATACTGTA TGGAATAACC TTACCTGTGG TAAGGATATG
148601 GAGGAGGAGG CTGTTTTAGA AGCTCTAAAA CGTGCCTACG CTGATGAGTT
148651 TATTTTAAAG CTCCCTAAAG GAGTCCATAG CGTGCTCGAA GAATCTGGGA
148701 AGAATCTCTC AGGAGGACAG CAGCAACGTT TGGCAATAGC ACGTGCTCTG
148751 TTGAAAAACG CCTCCATCTT AATTTTAGAT GAGGCAACGT CAGCTCTAGA
148801 TGCCATTAGT GAAAATTACA TTAAGAATAT CATTGGAGAG CTTAAAGGAC
148851 AGTGCACACA AATCATTATT GCCCACAAGC TGACCACTCT TGAACATGTA
148901 GATCGCGTGC TCTACATAGA AAATGGTCAA AAAATTGCCG AAGGCACAAA
148951 AGAAGAACTC TTACAGACGT GTCCTGAATT TTTAAAAATG TGGGAGCTCT
149001 CAGGGACTAA AGAATATAAC AGGGTCTTTG TTCCTGATCA CAAATTAGTC
149051 GCAAATCCTA CGGACATGGC AATAACAAC TTAGGTGGGAT CGCTCTCTCC
149101 ATGAGCTCAG GCAACAAC TACAAGTGTG TGAGTTAGCT TTTGTGATAC
149151 CTCCTCCAAT CTGCTGAAGG GACAGTCTCC TGGAACAGTA TAATCAGAAG
149201 TGATCTTGAG AAAAGAACAG GGGATGTGAT GTTCTGCTGC TTGTGAGGCT
149251 ATAGCATAGC CTTCCATATC TAGAAGTTTA AACGTCTTAT GAAACCCATA
149301 ATGGTACAAT ACTGGAGAGG TAACCAGAGA GCTTTTAGGT AGAGAATCCG
149351 GTAGAGCGTC AAAGATATAA GGGGGATCTT CAGAGAGAAC AGGAGGTGTA
149401 TCCGTAGTGA GGTTCGCAAT TTTCTCAATA GTGTAACATT GACCTAAAGG
149451 AATCTCGGGA GAACATGCCC CCACAAAACC TGGATTGATC CACAGATCGT
149501 AATCTGTATA TGCTTGCAA TAGCTTTGAA GAGCATTTAA AACGGCTGTA
149551 CTTCCCCAAA CATGGACAAT ATAGAGATCT AGATGGTAGT CAGTACAACG
149601 ATAACTATAG AGATGCTCGT TGATCTGTGT AAAATCAAGT TGTTCAATTA
149651 GAGGAGAAAT TTCTCTATAG TCTGCAACAA TGCAAAGGAT TTTTTTAGGT
149701 GTATTGACAG CATTCATTGG CCTTCCAGAG CATATGTAAA GCTTTTTTCC
149751 CAGTTTTAGA TAGTTGAAAG GTTTCTTTGT TGATATAGGT TCCTATGAAT

149801 CTATGAATCA CGGTCACGTT TTTATTTTTA GAGTATTCTA CTGCTTTTGC
 149851 TCCCGCAGTT ATAGGATCTT TCAGGGAGCA AATTAAAGAC TTTCTTAATG
 149901 CTGCTGTTAG AGCATCCACT GTAGCCATAG GAACATATTT CGCAATGGCT
 149951 AAACATCCTA AAGGAAGGGG AAAGATGGTC TTACGGCGCC ATAGCTCTCC
 150001 AAAGTCTGCC CGCAATGTCA ATTGGAGATC GTAGCTGAAG CGCTCTTCAT
 150051 GAATCAGAGC GCCTCCATCG ACTTTCCCTT GCAGTATCGC GGATAGAATT
 150101 TTGTCATAAG GCATGGGAAT GAGTTTTGCC TTGGGATAGT AAAGTTTACA
 150151 GAGAGCATGA GCGGTTGTCA TCTCTCCAGG AGTTGCCAAG GTATCTAGAG
 150201 AACATTCAGG ATCTAAGGAG AGGACGATAG GACCGCTGTT GTATCCTAAG
 150251 GTATTTCCCTA CGTCCATAAG ATTATAATAA TCAGAACTA GAGGGAAGAG
 150301 CGCTGCTGAC ATTTTCATTA GGGAGAGCCG TCGCTGCAGA GCTAGGGTAT
 150351 TCAAAGTTTC AATATCCGCA ATTGTTACCT GGTAAAGAAG AGGCCTGAAT
 150401 TGGGGTCTT TTAAGAAAGA ACGAAAAAGG AAAATATCAT TCGGGCAAGG
 150451 AGAAAAGGCA GCAGTCAGTA TCATGTCGGT TGATGTAATA GAGCTATAGC
 150501 GGCTTTGATG TCTTTATTTT CAGGCTTATC CAAAGCTCCT TGGTTTTCCA
 150551 GCCATTCGAA GTAAGACTTA GGAATATCCA CAAGAGGCTG CCCTTTGTAT
 150601 TTGCCAAAAG GCATTTTGAA GACTTTCGGG TGATAGCTCT GTTGCAGCAA
 150651 GTCGAGGACT TGCTGGGGCG GTAAATCACC GATTAAAGAA GTAAATACCT
 150701 TGTGCAATAT CACTACGTCA TCTAGAGCTC GGTGTGCTTG ATTTTCAGCA
 150751 AAACCGTAAA CTTGTCTTAG GTATTGTAAA TTATGTTTTG GTAGATCGGG
 150801 GCGATATTTT TGTGCCCAT TTAGAGAGTC TATTGTACGG TTTGTCAGAG
 150851 GCTCTAAGGA ATGTCTGCGA CATTCCCTTAC CGAGTAGGGG GAAATCAAAA
 150901 CCGTCATTAT TATGAGCCAC TAAGATGCTG TCCTCTCCGC AAAATTTCTT
 150951 AAATCCCTCG TAGGCTTCAG GAAATTTGGG AGCAGAAAGT ACCGCATCCG
 151001 TAGTGATTCC ATGAATTTTG GATGCCTCAT CAGGAATGGG AATTTCCGGA
 151051 TTCACATAAG TAAGAAAGGA CTCATCTGTG AACTATTGT AGGCAGCAAT
 151101 TTCTATAATG CGATCTCTTT CTATTTGTGT TCCTGTGGTC TCCGTATCAT

151151 AGAAAATAAG AACATCCATA GTTTGACTAC TCATTACGTT TTTCTTGTG
151201 CTCTTGAAGA GCCTGACGTC TTAGTTCATC CAAATTCATA TTCCCAGAAG
151251 AGATCAACCC AATAGCATGA GAAAACTAT CACAGACTAG CTTTATTGTA
151301 TCGATATATA TCCGTAATAG TGTGTCATGA ATTTCTCCGT TTAGGCAGGG
151351 CAACACAAGC CGATAAAATA TCAATCCCTG TTCTTCATCC ATGCCAAAGC
151401 CGGGAATATC AATGTCCCTA TTTAAGAGAT GGAGTAAACG AGCTGTTGAT
151451 GCCTTATGAG ATTCATGCAA TTGGTAGGGA AGGTAACAAA TCAACTGCAG
151501 TATTTCTCCC TCACTGCGGA TTACAAAAA TAAAGGGAGT TCATTGCCAT
151551 TAGCTTGAAT GTTAATGTAA GTAAGACCGC TTTCTCTTTC TAAGAAAGGT
151601 TCTTCATCCG AACTTTTAAG AAATTTTGTG AGATTATTTT GATTTAATGT
151651 CCATGTCGTC ATTTAGGAAA TACTCCAAGT TGTTCCTAGA GCCTGCATCA
151701 TTGCTGGCTG ATAATACTAG ATCTAATCTT GATTCGTCTG TTGTTTTTTT
151751 TGTATCTTTT GTATTGCTTT ACTGAGGAAA TCAGAGGTTG CAGAGGAAGA
151801 AGACTGCTTA TTATTTTCTT GAAGTAGATG ATTGATATCA GTGTCGGGAG
151851 GAAACTCTGT CAGTAGCGTA ATTTTTTCTG GTGTTTCTGC AATCACAAGA
151901 GTTTTATTCA CAACTCGAAT GAGGTAAATA GAAGTTTTCG GCGTTAGGGA
151951 ACGTCGTTCT AGGATTTTGA TTTGAGACGA GCCTCCAAA CCGTGACTTC
152001 TTGATCTCAC AAACTTTTTA AACGCCCAA CTCCAAAGCC AAAAATTGTT
152051 AAAAGTAGAA TCAAAGATCC TAGCATTTTA AACATTTCTA ATTTCATGCT
152101 TCCTGGGAAC ATTTTCATGTA CAGAAATGGG CTCTTGGATC GTTTCTGCAA
152151 GAGCAAGCTC ATCAGAAAGC TTAAAACTA AAGAAAAAAG ATTAAAAAAC
152201 ATGTGTAAGA CCCGCGATCA TCTCTATAAA ATTATAGTGG TAGCCCGATT
152251 TGTATCCAAC TACATACAAG TAATAATGAA GTATAGTTTT AGTCGATGCT
152301 ATATAAATTA TAGTACAATG ATTTCCAAGT ACAGATTAAA CCGTAATCAT
152351 GTATATTCCT GCAAGTACCG TCTAGAGAGC TCCCCTAGAT GATTTGGGTA
152401 ATTCACAGAC TCCTTTATAC CCTTCTAGGG TGTGCTCGTT CCACAGAGCC
152451 CAAGCTCTTG TCTTTCATAG ACAAACGAC AGCAGTCTGT CGGTGGATGC

152501 AAAGAATTTT TTCAAGTTGC CTGAGAAATT CCTTGGCAGC GCTTAATAAA
 152551 ACTATGGTGA TGCTATGGAA AAGTTACTAG TGA CTGATAT TGACGGTACA
 152601 ATTACCCATC AATCTCATCA TTTAGATAAA AAGGTGTATG AGCGGCTCTA
 152651 TGC GCTGCAC CAAGCTGGTT GGAAGTTGTT TTTCTTGACG GGAAGGTATT
 152701 ATAAATATGC TGCACGCTTG TTTTCTGATT TTGATGCTCC ATATTTATTA
 152751 GGATGCCAAA ACGGCGCTTC TGTATGGTCT TCAACATCAT CAAATCTTCT
 152801 CTATTCTAAA AGTTTACCCT CAGATTTATT ATGTATTTTA CAAGATTGTA
 152851 TGGAGGGGGC AACGGCTCTT TTTTCCGTGG AATCAGGAGC TCCTTACGGG
 152901 GATCACTACT ATCGCTTTTC ACCGACTCCT ATAGCTCAAG ATTTACACGA
 152951 ATATGTAGAT CCTAGGTACT TTCCTAATGC TAAGGAAAGA GAGATCCTAT
 153001 TTGAAACGCG CTCTTTAAAA GACGACTATG CTTTTCCTAG TTTTGCTGCA
 153051 GCAAAAGTCT TTG GACTGCG AGATGAGGTC ATCAGAATTC AAAAGGAGCT
 153101 GGAACGCCAA GAAGCACTGA CTTCA GTCGC GACGATGACG TTAATGCGCT
 153151 GGCCCTTTGA CTTTCGCTAT GCCATCTTGT TTTTAACAGA TAAAAGCGTC
 153201 TCTAAAGGCA AAGCCTTAGA TCGTGTTGTC AATATACTTT ATGATGGAAA
 153251 GAAACCCTTT GTCATGGCTT CAGGAGATGA TGCTAATGAT CTCGATCTTA
 153301 TTGAGAGAGG AGATTTTAAA ATTGTGATGA GTTCCGCACC TGAAGAGATG
 153351 CACGTTCATG CGGACTTTCT AGCTCCCCCA GCAGATAAGA ATGGCATTCT
 153401 TTCAGCTTGG GAAGCTGGTG TCCGCTATTA TGACGACCTT ATGAGTCTTT
 153451 AGGGAACATC TCAGGACCAA TTCCCATCAC ATTGGCTCCG TGATCTACGT
 153501 ATAAGGTCTC ACCAGTAATT GCTGAAGCTA GAGGTGATGC TAAGAAAGCT
 153551 GCAACGGCAC CCACCTGCTC GGCATT CATA GCCTCGGGAA TAGGCGCCCA
 153601 CTCTTGGTAA TAGTCTACCA TTCTTTCAAT AAAACCAATT GCTTTTCCAG
 153651 CTCGCTTGC TAAAGGTCCT GCAGAGATGG TATTGACACG TATGCCCCAA
 153701 CGGCGTCCCG CTTCCCAAGC AAGAGTTT TG GTG TCACTTT CCAAAGCTGC
 153751 TTTTGCCGAA CTCATGCCCC CTCCGTATCC AGGAACAGCG CGCATAGAAG
 153801 CCAAATAGGT GAGCGATATT GTCGATCCAC CACGGTTCAT GATACTTCCA

153851 AAGTGAGAGA GAAGGCTAAC AAAAGAATAA CTAGAGGCAC TGAGAGCCGC
 153901 TAAGTAACCT TTTCTTGATG TTTCTAATAG AGACTTAGAA ATTTTCAGGAC
 153951 TATTTGCCAG CGAGTGGACA AGAATGTCAA TATGACCAAA ATCTTTTTTT
 154001 ACCTGTTCTG CGACTTCTGA TATCGTGAAT CCCGTAATGC CCTTGTAACG
 154051 TTTATTTTCA GCAATATCTT CAGGAACATC TTCAGGGCTA TCAAAACTTG
 154101 CGTCCATGGG ATAGATCTTA GCAATCTCTA AGAGAGTGCC ATTCGATAAT
 154151 TTTCTAGATT CATTGAATTT TCCTAATTCC CAAGACTGAG AGAAAATTTT
 154201 GTAAATCGGT ACCCATGTTT CTACAATAAT CGTAGCTCCT GCTTCTGCAA
 154251 GAAGTTTAGC AATACCCAG CCATATCCTT GGTCATCACC AATGCCCCGA
 154301 ACAAATGCTA CCTTTCCTGT TAGATCAATC TTTAGCATGA ATCCGCCTTA
 154351 TACTTTTGAA GCTTATTGGA AGGAGAGTAA CAAATCTTTC GATTATTAAG
 154401 AAAACCTTTT GGTGCCTCAA CAGGGGAGAT CCTGCCTCCA ATGTAAATAG
 154451 AAACGTAAAT TCTTTAAATT TTTTCTTTA CATATTTTAT AGAATATCCA
 154501 AACTTCTCAC TCCCGCGTAC TGCTAAAAAA ATTTTCAAAA GAATTTACGA
 154551 TCCGAACCTA TCGTAGTTTG GGTTCCTACTG ATTACTTAGG AGGTTGTTTG
 154601 ACGAATCCTT TAGGGAAATT CCCCTCACCA CAGAATCCAC AGGTTGTTAC
 154651 GATAGCGCCT TCTTCCACAA CACCACAAGC AGTCTCATCT GCAGTTCAAG
 154701 GTTTTCTTCA AACTGGAGGA GCTGCCTCCT CTACAGCGAC AACTACTACC
 154751 GCATCCGGAG CCTCTGCATT AGGACTTTCA CCTGATCAAG TGCAAGCGTT
 154801 GCTTACTAAT TTATTAAATG TGGGACAACC ATCAGTGGGA CAACCATCAA
 154851 CTTCAGCAGG AACTTCGGGA GCCTCCTCTT CCAGTGCAAG TATGCAGCAA
 154901 CAGCTTTTGC AACTTATCTT AGACAAGACA ACAGGAAGTG GCGGATCGTC
 154951 CGTGAGTTCA GAGCAATTAC AGCAACTCCT TAGCTTGGTG AGCCAGATGA
 155001 CTACGTCTCA AGGAGGAAGT GGTGGAAGTC AGGCAGGACA GGCCGCTTCG
 155051 GTACTGTTGA ATTTGTTATC GGCAACAGGA TCTGCAGCAG CAAATCCTTT
 155101 AGGGACAGCT GCATCGTTGG CACAGATCAT TTATGCAGCA GTAACAAGTC
 155151 CTGGAGCAAA GAAAAC TAGC GAATTTTGTT ATAATTATTG TGGAGAGACC

155201 TGCCAAGGCA ACTGCGGTTG TCCTACCTGT GGCTGTCCAG ACGGACAGTG
155251 CGGTTGTGGA GGATTTGGCC GTTTTTCTG TGGTGTATGG AAAAATTGTT
155301 GCGGGATAGG AGAGGGATCC CAAGAACCCG CAATCCCTTT ATAAGAACTC
155351 GAGGCTTTAG AACAGAAATA TGGCAAGGCT GTTCTTTTAA TTGCGTTAAG
155401 TGAGCTTGGC ATTGATACCA TGAGCTTATT ATCAGGACAT CGACTCGAGG
155451 GATTTCCTCC AATCGCGGAG GTCATGGCTG CATGTGACCG GTGTTCTATG
155501 GACTTTTGTG AGATCTTGAA GTCTCAAAGC ATGGATCTGT GGGCGGATGC
155551 GCGGAGTTGT GTGGATGGTT TATTACAGGA TCCTTTTGG AGTACAGCAA
155601 TTGCCTCAGG GATTGCTAAG TCTTCTCTTC AGGAAACGGA ATTCGAGTGT
155651 GAAAGCAAAG TGATGGTTCT TTCTTCATGG GGAGAGCAAG GAGCACAGGT
155701 TTGTAGCCCT TTTAACCTAG AGAGGATATG TATGTCTTTC CCATCACTTA
155751 AGGTCTTCTC CCTTAAAAAG AACGGGTGCG AGAACATGGG AATCCAGTTG
155801 TCTGCATCCT GCATGAATCT ATTAATGTCT ATTTTCTTTG TAGCTACCAA
155851 TGGAGGAAGC ACTCCGATTT GGATCACCAA AGAAAATCTG ATGGCGTTAG
155901 TTGCTTTGGT TTTATCTCAC TATCAATGTT ATTTTGTCCC AGCCACAGGA
155951 GATCCCCAAC GTGGCAACAT TTTAGGTAAT CCAGAAGTCA ATGCTATTTT
156001 GGCTCGGGGG ATGGGCATGC GTGTGATCT GGAAAGGAAG CGAGGGGGAG
156051 AATCTTCCTC GTCACGCTAT TTAGAATTAG CTGCACGATG TTTTGAGAAT
156101 TCTCTTACGA AAACAAGTTT GTTAAGCGAT GCTAACAATG TTCAAGAAAG
156151 AGATAAGTGC CTAACACAGA TGTCAACTTC ATTGATGCAT ACGGCGGGAC
156201 TAAATTTACA ACGCCCCCCT GTACCCACAC CTTCTGGAGT CACGGCACAT
156251 CCGCAACCTC AACCAGATCC TGTGGTTACG TCTCAACCTT CTTTATTAGG
156301 TGCTAGAGAG CGTTCCTCTG TGTCTTCTAG AGGGCGTTTT CCTGTAGTTT
156351 TACCTTTAAG TGTGATTTCT CCTAGGTCGC ACCCCGGAAG GGTAGAAAGG
156401 CGGGATTTAG AAGATGAAGA AGAGGAGGTT ATGTTTTGAA GCAGTGATAA
156451 CGACTCCAAT TACAGTTTTA TGAATCTCTA ATTGTAAAGT TCTAGGGGTT
156501 TTTCTTGAAG TAAGTGCCGA GCACATTCTC TAGGATCTTC GGTGATGAC

156551 GCACAAATTT TTAGGGGAAG ATTTGTGAAT GGGGAGATAA ATTCTAGGGA
 156601 GTGAGCATGG AGGAGAGGGC GGAAGATCTG GGGAGGCTGT TCTTTAGGTC
 156651 CGTAGTCGAC ATCTCCGACA ATAGGATGAC CCAGCAATCC CATTTGTAAG
 156701 CGGATTTGAT GGGTTCTCCC TGTGATGGGC CTGCGGCTCC AAAAATCACA
 156751 GCTCCACACC TCCGGTATAC GGGGGCCGTA TAAGATTTTA CGGTTCCAAA
 156801 TTTTTTTTTT GGATGACCAA AAACGAAAGC TATGTATTGT TTATGGATTT
 156851 TTCTTTGCTT GAACAATTC ATGAGCTCAG TAGCCGCTTG TTTAGACTTT
 156901 CCCATGAGAA GACACCCAGA GGTGCCTTTG TCTAACCTAT GCACAGTAAA
 156951 AAACCGTGTC ATGTGTGCCA TTTGTTCACT AGTAAGATGG GGAGGTTTTT
 157001 CGTAGATAAT GCTATAGTCA TCCTCCCAGA GGATGCTAGG TTGTTGTTTT
 157051 GTTGAGGGGA TCAGAGATAG GGAAACACGG TCGCCAGGTT GTACCTTGTA
 157101 GGATTCAAAT CTTTCTATGA ACCCGTTCAC TCGACATCGA TGTTGGCGAA
 157151 TAGACGCCAA GATTTCTTGC TTGCTATGAT TAGGCAGTTG AGATCTAAGA
 157201 AAAGAAGATA ATCTTGAGAC TTGTGTGGCA AGCCAGGAAA AATTTTCCAT
 157251 AAAATATTGT AAAGCAGCCC TTTTATCATT TGATAATTGC ATAAAATTTT
 157301 AAGAGATTTT GTATGACAAA GATAGCTTTT TCTGAAAAGG CAAAGAATTT
 157351 TCCTGTAGAG GCATTAAAAA AATGGTTTGA AAAAAATAAA CGATCTCTTC
 157401 CTTGGAGAGA TAACCCGACT CCCTATAGTG TGTGGGTTTC CGAAGTTATG
 157451 CTACAGCAAA CGCGAGCTGA AGTTGTTATA GATTATTTTA ATCAGTGGAT
 157501 GGAGAGATTT CCTACCATAG AGTCTTTAGC TGCAGCAAAA GAAGAAGATG
 157551 TCATTAAGTT ATGGGAGGGA TTGGGTTATT ATTCTCGAGC GCGCCATCTT
 157601 TTAGAGGGAG CTCGCATGGT TATGGAGGAG TTTCATGGAA AGATCCCTGA
 157651 TGATGCCATT TCCTTAGCTC AAATTCGTGG AGTTGGTCCT TATACGGTTC
 157701 ATGCTATTCT AGCCTTTGCT TTTAAGAGGC GTGCTGCTGC TGTGGATGGC
 157751 AATGTCTTGC GTGTTCTTAG CCGGATATTT TTGATAGAAA CTTCTATAGA
 157801 CTTAGAATCA ACTCGTACTT GGGTTTCTAG GATTGCTCAA GCGCTTCTTC
 157851 CTCATAAGAG TCCCGAGGTT ATAGCTGAGG CTCTGATAGA GTTGGGAGCT

157901 TGTATCTGTA AAAAAGTTCC TCAATGTCAT CGTTGTCCTG TCCGTCAAGC
157951 ATGTGGAGCT TGGAGGGAGA ACAAACAGTT CGTATTGCCG GTACGTCATG
158001 CCAGAAAAAA GGTCATCTTT TTGCATCGTT TGGTAGCGAT TGTATTGTAC
158051 GATGGCTCTT TGGTTGTCGA GAAGAGACGT CCTAAAGAAA TGATGGCAGG
158101 CTTATATGAA TTTCTTATA TTGAAGTTGA ACCAGAGGAA GGTCTTCAAG
158151 ATATAGAAGG ATTTACTAAG AAGATGGAGC TTTCTTTAGA AAGCCCTTTG
158201 GAATTCTTAG GTAACCTTAA AGAACAGCGG CATGCGTTTA CTAATCATAA
158251 GGTTCATTTG TGTCTATAA TTTTAAAGC CACTTCTCTG CCTCAGTTCCG
158301 GGAATTGCA TCTTTTGAGT GATATAGATC ACTTAGCTTT TTCTTCAGGA
158351 CACAAAAAGA TTAAAGATGC TTTGCTAATC TACCTCGGGG ATGTCAGGTC
158401 TAGAGAATCA ATAGGAGTAT AGATGCGAGA TCACGCTTTT TCTAAATTGA
158451 TAGGGACTGT CCGTGCCATG GTAGTTGAAG GACGTTGTCC TTGGTCACTT
158501 CAGCAATCCC TAGTCTCTAT GGTAGAGCAT ATTCTTGAG AGTGTCAGGA
158551 ATTTACAGAG GCCGTCTTAC AAGGTAAGAC GGTACAAGAG GTTGGTTCGG
158601 AAGCCGGGGA TGTCTTAACT TTAGTTCTAA TTTTATGTTT TCTGTTAGAA
158651 CGAGAGGGCG TACTTGCTTC CGAAGACGTT GCCAATGAGG CTATGGAAAA
158701 ATTGCGTCGC CGTGCTCCTT ATATATTCGC TGAAGATTAC AAGCCGGTCT
158751 CGATTGAAGA GGCCGATCGC CTTTGGGAGC TTGCTAAGCA CCGAGAGAAA
158801 AATGAATCTA CATAGTTGAA GTTTTGGTCT ATTTTAAAGC ATATGGTGCT
158851 TTTGAAAAAA CAGAATATAT GCTATCAAAG AAGGGTAAGT TGGGGGCCTT
158901 TTAAGAGAAG GAACCTGCGA ATCGGGTCAG GACTGGAAGG TAGCAGCCCT
158951 AAGGAGAGTT TTCTTTTGCT AAAAGAATGT TCTCCAAGT ACTCTTTTAA
159001 CTTTATTCCC AAAAATAGCA ATGAGGTGAG GTTAAACAAC CCGTGCAGTG
159051 CAATGGGAGA AAGAATGTGC CGATCTTTTT CATATAGAAA CCCTGCAGAT
159101 AAGGAAAAAA CAAAGAGCAC GGGGACAAAG ACCCAACTTC CTAAAGAGTG
159151 TTCAATGTGA ATGAAAGAGA AAATAATAGA AGAGCATAGT ACCGCAGCTA
159201 TGCGCGTCAT TTTGTTTTTC AAGAATGTCT GTAGAATTCC TCTAAAAAAT

159251 ACCTCTTCTC CAAATGGAGT GAGGACGCCT AAATTTAGAA TCATGCTAAT
 159301 GTAGTGTCCCT GTTATAGGCA GAGAGTTCTG AACTTCTTGA GTGACTTCTT
 159351 GTGTGTGAAT CTCTTGCGTA GGAAGAACCA AAGTTAAAAA TTTACTCATC
 159401 ATAATCCCAA TCAGTTGTGT TACTGGGATG ATGATGATCC ACATTCTGAT
 159451 GGCAGATCCT AGAGCACGCC ATGAAGTTTT AACCGGTCTT TCTCCAGAGA
 159501 AAAGTATAGC ACGTGTGATA TCCTTGGGGA GAAAAAGCAG GTAGAACAGA
 159551 AATGCAAAGG CAAGGCTAAT TCCTGTCATG GTGGAAAGTA ATTCCGCAGT
 159601 TTGTGAGCTC ACGCTAAGAG CTACAAGGGA AGAAAAACA AGAAGAGCAC
 159651 CACCAAATAA AACTTGGCGG AGCTTTAAAG GTGTTTTCCC AGAGGGTGCT
 159701 GGCCAGATAA AGAAGTTTTT GGAAGCTAGA GCAGCGACGC CAAGGGACAA
 159751 GAGAAGAATA AACTTGGACA TTTCCTTAGA CTACGAGTAG TTAGCACAAA
 159801 CATAGCCCTC AACTCTGGCA ACAACTTCGC GGAAAAGACG GCTATGCATT
 159851 AAGCCCATGG GCGTTGAATC AAAATGTTTT GAGTTCCAGC CATAGCGATG
 159901 ATAATCATTG ACTAGGGTAG TTAAAGGCTG GCTGCATTCTG ATAATCTCTT
 159951 GATAAATGAG AGCTATTTTA TGATGACGGA TATCAAAAAC GCGAACACGT
 160001 ACAGACGCTG TTACAGAATC GACACCTGCT TCTTTCCCTG TCTTTTGTTT
 160051 TAACAGTTCT GTAGCAACAA TGAATTCTGC AGGAAGAAAT TGCTCAATAA
 160101 TTGTTTCGGG TAGACGATTC GCAATCGGAG CATAGAACTG AGAGACTGTC
 160151 TGAGGTGAAG CATTGTGCTT GATCAGGAAG ACCTTTTCCG AAGCATAAAA
 160201 CCTTTTGCTG ATCTCTTCAG TAAATTCTCC TTGGAGGTTT CAAGGTAAAG
 160251 GTTCAAGACT CTTTCCTGGG CGATGAAATA CAGGAAGCAT CGCAATCACA
 160301 CCTTTAGTTT TGCTCCCTGA AGTGTATAGC TTAGGATGAT AACTTCCTGA
 160351 AGAGCCTAAG TGAGTGCAGC TGGATAGGGT TGGGGATAGA AGTCCTAAAG
 160401 ATGCCAATAA TACCAACATT TTTCGCATAG TCACTGTCCT TAAATTGCTT
 160451 ATTTTGCAAA AGATTCTAGC CCTGGGAAAG TTTTACTTTT TAAGATCAAT
 160501 ACTTTCGCAA TTGAGAGATT TTCCATTTAA AACTCTCATT AGCTTATATC
 160551 AAAGAAAAAA ATAAAAACAA GCAAAGAGAC CGTCTCAGTT TTAGTTTAGA

160601 AACTCAAGGT TGAGAAAGGG ATTCTGACCA AAGTTGTGAG GGAAC TTTGG
 160651 TAACTTTTTC TTTAGGAATC AATGTGCACC CTGGGAGGAA TACAACTGGA
 160701 GCTGAGATCA CAGAAAAAAT AAACCACTTC ATGACTACCT CTGCAAAAAG
 160751 AATACTACTA TTTTCTATTT GCATAGGCAG TTCTTCGATT TATAGCAATT
 160801 TTTACTTTAT CATCATAAAA ACTATGATGA AAAGGTTCTT AGTGAAC TTC
 160851 TAAGGAACAC CATGAGTTTG TGATTAAAAC TCATGGTGCG TAGTTGAACC
 160901 CTTATTGAGG AGGGACGCAA CCACTAAGAG CTAACAATAC TAAACAAGAT
 160951 AATAGTAAAG AGAATGCTTT TTTCATTATT CATCCTTAGG TTAATTGACC
 161001 TTTCCAGCCT AGCTCTAGGG CGATTATTTA TCAAATTTTT CTTTGTAATT
 161051 AATGATCATG CGACCATTAA TTTAGCGATA AATTATGATT TCGTCAGGAA
 161101 AATTCAATTC TTTATAATAA TGATATGAAA TTAGAGAATG TCTATAGGGG
 161151 CGGACTCTAT TTGTGATCCA GGATCTCTTT AGGAGCACTT TGTGGATTTT
 161201 GATTATTTTG GTCTGAGTGA TATTGGTAGG GTGCGCGCTA GAAATGAAGA
 161251 TTTTTGGCAG GTAAACCTCA TGTCTCAAGT GGTGCTATT GCTGACGGTG
 161301 TTGGGGGGCG TCTTGGTGGA GACATTGCTT CTCAAGAGGC AGTGA CTAGC
 161351 CTTATGGAGC TGATTGATGA GCAACAGTCA AAATTGATGG GGTATGGGGA
 161401 TGACCAGTAT AAGGAGACTT TAAAAAAGAT CCTTTTAGAG GTCAATGGTG
 161451 TGGTCTATGA ACACGGCCAA ATGGAAGAGC ATCTCCAGGG TATGGGAACC
 161501 ACTCTTAGCT TCATCCAATT CCGGAAGGAT AGGGCATGGC TATTT CATGT
 161551 GGGAGATAGT CGAATTTATC GTATTCGTGA GGGAGAACTG CGCCGCCTTA
 161601 CCGAAGACCA TTCTTTAGAA AATCAATTAA AAAATCGTTA TGGGCTTCCT
 161651 AAACAATCAG ATAAGGTGTA TTCTTATCGC CATATTCTGA CTAATGTTTT
 161701 GGGGAAGTCGT CCCTATGTCA TGCCTGACAT TCGGAATCTT CCTTGTGAAA
 161751 AGGAAGATTT G TACTGCCTC TGTTCCGATG GATTGACAAA CATGGTTCCA
 161801 GATATCGATA TTCGTGATAT CTTGAACCAG CCCGCCACCC TAGAAGAACG
 161851 GGGGAATGCA TTAATTTCTC TAGCCAATAC TCGTGGAGGC GATGACAACG
 161901 C TACTGTCGT ATTAGTCCGA ATACAATAGT TCCTTTGCTA AGGATAGTAT

161951 TCCATGATCT ATTTGGATAA CAATGCGATG ACACCCCCAG AGAGGGGACT
162001 TTTGGAATTT CTCCAAAAA CCTTCCTTAT AGAAGGGACG TACGCGAATC
162051 CTTCGAGCGT CCATCAATTA GGTAAAAAAT CTCGTCAACT GGTTCCTAGAA
162101 GCTTCACACT GGATGCAAAA GGTCTTTTCG TTTCAGGGCC GTGTCCTCTA
162151 TACCTCAGGG GCTACTGAGA GTTTAAATTT AGCAATAGCA AGCCTCCCTA
162201 AAGACAGTCA TGTTATCACC TCAGGTAGCG AACACCCCGC CATCTTAGAG
162251 CCTTTAAAC ATTCTCGCT TTCCGTTTCT TATTTAAATC CCGAAGAAGG
162301 GAGATGTGTT CTTACTATAG AGCAGATTGA AAGAGCTGTG ACTCCTAAAA
162351 CTTCAGCAAT CATCTTAGGT TGGGTCAATA GTGAGACTGG TGCCAAAGCT
162401 GATATAGCTG CTATAGCCCA CTTGCGCAA GAACGACAAT TGCAATTTAT
162451 TGTGGATGCG ACTGCAAATG TAGGTAAGGA GAGGATAGTT CTTCCCTCTG
162501 GTGTCACTAT GGCAGCATTC AGTGGACATA AATTTTCATGC ACTCTCTGGA
162551 ATCGGAGCTC TTCTGGTCTC TCCAGGAGTC AAACACATC CTCAGCTGTG
162601 GGGAGGAGGT CAGCAAGGAG GGCTGCGCGC AGGCACAGAA AATCTTTGGG
162651 GAATCGCCTC TCTGCTTTAT ATTTTCAAAT ACCTAGATCT TCATCAAGAG
162701 CGTATCTCTC AGGAAATTCT TACCCATAGA AATGGTTTTG AAAAGGCAAT
162751 CAAAGCACGC ATTCCTGATG TCCATATTCA TTGTGCGGAT CAACCACGGG
162801 CAAACAACGT CTCAGCAATT GCTTTCCCTC CGTTGGAAGG TGAGGTATTG
162851 CAAATCGCCT TAGATATAGA AGGAGTGGCT TGTGGTTATG GATCCGCATG
162901 CTCTTCAGGT GCTACCGCAC CCTTTAAATC TCTTGTCAGC ATGGGTGTTG
162951 ATGAAGAGTT GACCCTGGCA ACACTCAGGT TTTCTTTTAG CCATCTTCTC
163001 TTGCAAGAAG ATGTTGAAAG AGCCGTTGGA ATTATAGAAA AAGTCGTAGA
163051 ACGTTTGAAA AATTCCTAAG TCTTAAAAGA GAACATGTTT CTAAGCTGAA
163101 AGAACACTCC TGACTCTTAT TGCAGAATCT ATGAGAGTAA GTTTTAAATC
163151 GATACGGTTT TTATCCCAGA TAAAGACATC TCTTTAACTT CTAAAAGCAA
163201 GTTGTGATA ATTACAGAAG TTCCTACTTC TGCAGGACTG TCTAGCAGTT
163251 GCAATACCAA TTGGGCTAGG GTTCTACAG GATATTGCGG AAATTGAATA

163301 TCGAGTTCTT TTTGCAGATC TTTTATGCGA GAGTTGCCAG GAAACGTTCT
 163351 TTCAATAACA GAGATGGTCT TGGGTTTTAA ATGAGCAATG TTTGTAGTGT
 163401 TGAATAAGAT TTTGAAAATT GCATTTAAAC TAAGAATACC TATAGGTTCA
 163451 CCAGAAGCAT TGAGGACAAC AGCAACACTC GAACGGTTGT CTCGAAACTC
 163501 TTTGAGGATA CGAATAAGTT TTGATTTTGC AGTGATAAAC CAAGGCGAGT
 163551 GTAGATTATT GATTAGGGGT TCATCAAGAG CTTTATTGAC AAAGTCTTTA
 163601 GGATGGGCAA TCCCAATAAC GTTTTTTCGG GCCTTGTGAT AGACAGGAAT
 163651 AAAGTTGATA TCTGTATTTT TTATAGTCCG GCAAAAATCT TTAACATTTG
 163701 CAGAAGAAGG AAGCATGGTA ACCTGTTCTA AAGGTTGGCA TACCTGATCT
 163751 GCACAAGTCG CACTTAAAGA GAAAATATTT GTAGCAATTG TATTGAAATC
 163801 TTGTTCTTCA TGGTGAGTCT CTAAAGCTTT TTGGAACCTG TCTCTACTTA
 163851 ATGTAGAGTT CAATTTTTCT TTCCTAATAT TTAGAAGATA GTAAAGACCC
 163901 TCAGTGAGAC TTCCTATGAG CTGAATCAGA GGATAGAAAA TATAGTGGGA
 163951 ATAATAGAGA ATCGGTGCTC CCCAAAGTGC TAATTTTTCA GGAATCTTCC
 164001 GTGATATTGT TAGAGGTAGA AGTTCTGCAA AAATCACAAC TATAAAAATT
 164051 TGAGTGAAAG GAGCGTAATC TGGAGTGATT CCTAAAGCTC GATAGCAATT
 164101 TCTTGAGGAC TCAGACCCGA CTGTAGAGC GATATTCACT CCTAACATCA
 164151 CCGTTCCAAA TAAACGATAG GGGCGGCGAA TCAGGAAATT AATGTAGCGA
 164201 GCTTTCTTAT GATCTTTAGT CAGATAGTAT TGCAATCGTA CACGGTTAAA
 164251 TGACACGCAG GCCATTTCCA TCATCGAATA GAATCCTTGT AAGACAATAC
 164301 AGATAATGTT GACTCCTATC CAAAAGAGAG CAGAATTAGT CATACAATTT
 164351 CCTTATATAC ACACGGCGAA TCGGATTCGG AGCAGCGTCT AATACCTGGA
 164401 AAAGCAAGTT ATTCCAAGAG AGTTTCATTC CTGTTGTCGG AATCGTTCGG
 164451 ATTTGCTCTA TTAACCAGCC TCCTATAGTC GCAATATTAT TGTTGTCGG
 164501 TAGGTTGATA TCGAAGATCT CACTAAACTC ACGGAGTTCT AAAGTTCCTG
 164551 AGGCAATAAT AACATCAGCT CCTGAGGTGG TATAGAGTAT TTTATTATCT
 164601 CTCTGGTCTA CAATTTCTCC AGCAACAATT TCAAAGAGGT CTTCTTGAGT

164651 GATCAATCCT TCAATAGATC CGTATTCATC AATGATCATC CCTAGGGTTT
 164701 CGTCTTCAGC TGCCATCTGA CATAAAGCCA TTTTTCGAGA GATGGTTTCT
 164751 GGCATATAAT ACGGTTTTTT CAGCAAGGGG AGGAGATCAT CCGAAGATTG
 164801 CAGTGGCTTG TCATGTAAAA GAAGAGAGCG CGCTGTGCAA ATGCCCAGAA
 164851 GGTTTTGGAG GTTATCGTTA CATATAGGAA CTCGTGAGCA ATGCTGTTTA
 164901 GAAAATAAAA GATAGAGGTT CTCTAAAGGG GTTTGGATAT CATAAAATAA
 164951 AATATCCTGG CGTGGCTGCA TACGCTCTTT AACACTACAA TCACTAAGAG
 165001 AAAGATAACC ATAGAGTAAA CGGCTTTCTT CTTGATTGAC TACGCCGAAA
 165051 TCCTTACAAC TTTGCAATAC TTCCTTCAGC TCTTGGGGTT GGATGATATC
 165101 AATCTGTTGC TTCGATAAAA TCCATTGGAC CACATAATTA ATTCTACGA
 165151 TACCCAGTG GAGTAGGGGT TTGAAGATTT TAGTAACACA AAGAATAAGA
 165201 GGGGCTACGG AACTAGCAAT CTGTGTATTA AAAGGAAGAG CTACTGCTTT
 165251 AGGGAGAATC TCACCTAAGA TCAAAGTAAT TGCTAAAGGA AGACCTACAG
 165301 TAAACCACCA CGAAGCTGCA TCTCCAAATA GAATGGCAAA ACAGTTTTGA
 165351 ATAGCAATAT TCAGTCCGAT ATCACAAAAA ATTAAGGTGA TGAGCAGGTG
 165401 GTGGGGATGT AGAAGAAGGG TAGCTACTCG CTGCTGTTTC TTAGATTTAG
 165451 AGCGCTTATA GTGCGAGATC AACTCGTAG GCAAAGAAAA CAAAGCAATT
 165501 TGAGATAACG AAATGAATCC CGAGCATAAA GTAAACAGA TAATGAAGAA
 165551 CATTAACATG GTAGGAATCA TGGTCTCTTT TCAGTCCTTA TTTTCTGATT
 165601 GTTGCTTTGG GGAGACACAG AGTTTCTTAT AGCCTTTAGG AATGAGACCC
 165651 AGACGCTGGA TTAAAGCAGC TTCTATAGCA GCGGAGTCTT GCCAGTGTTG
 165701 CAGATGTAAT TGGAGTTGAC GCTGCTTTTC TTGAGCAGAA AGAATGTCTT
 165751 GGCATAAAGA AGAGACCTTG CTTTGTAAGC GTAGCTCTTC TGTACGTAAC
 165801 TCCTGGATAG CACGATCATA AACAAAGCCT CCAATTAAGA TGCTAAAGAT
 165851 CACCCACCAG GATTTGATCA TCACTTCTTC TAGTAATCTA AAACCCAGT
 165901 TTTTTTTTCT TACTGAAACT TTAGACACAA GGTACGGTGA TGCCTTGTTG
 165951 CTTTTGATAC TTTCCTTTTC TGTCTGCATA AGAAACCTCG CAGGTCGTAC

166001 TAGACTCAAA GAATAAGACC TGGGCAATCC CTCATTAGC GTAAATTTTC
166051 GCTGGCAATG GCGTAGTGTT AGAAATTTCT ATAGTCACAT GCCCTTCCCA
166101 TTCAGGCTCA AAAGGTGTGA CATTACGAT AATTCCACAG CGTGCATATG
166151 TAGACTTTCC TATACACATT GTTAAGACAT TTCTAGGAAT TCGGAAATAC
166201 TCAACGCTAC GAGCTAGAGC AAAAGAATTT GGAGGAACAA TACAGACGTC
166251 ATCAGTAATA GAGATGAAGA TATCCTCAGT AAAGCATTTT GGATCAACAA
166301 CAGAGTTATA GACATTGGTG AACACTTTGA ATTCTCGAGA TAGGCGGAGG
166351 TCGTAACCAT AACTCGATAG GCCGTAAGTT ATAAGTTTTT CGCCTGTCTC
166401 CTCATTTACG TTCACTTGGC CATTAACAAA GGGATGGATC ATATCGGCAT
166451 TTAGGGCCAT CTCTCGTATC CACTTATCTT CTTTTATGCT CATTTAGAAA
166501 CCTTAACAGT TTGAAATTGC TTTCTTAATG ATATTCTGTT TTTCAATTTA
166551 CTGGTTTTTG GGGGGAAGTT TTCTAAGTAT AAGATAGACT TTGATTATCT
166601 CTTGAAAAGA CCAGTTGTAT AAACAAGAAA AGCCTATCCC AAAGGCTACA
166651 ATTTTATCAC GAAACCTAGA GGTAATGTTA GATAATCCTA AGGGAAAAAG
166701 GCAAACCTTA TTTTATAGGA GAACTTCAGG TAGGTCTGCT CTTTACTCTT
166751 ATAGTAGAAG AATCTTGGTT CTCTTGAATG CATTCATGCG AGGACCTTGA
166801 TAAGAACTTC TTGGATTCAT AAAAAGATTA ACATCTCCTT ATTGATAAGC
166851 TAGAGAATTT TTAACACAA CTTCTCAGTG GAAAATGTTT TTAAAAATAG
166901 TTCGCCATCT TTAATTTATC TGTTTTAAGA CAAAAGAAAT CTAGATCACC
166951 ACAGGAAGTT TAAATCATAA AATGAAAATG ATGGAGAGGT TCTAGTGCTC
167001 GTACTTTGGC CCTGCTCTCC TTGATAGAAA GAAGAGGTCC ATAGTGACT
167051 TCTATATAGT ATCTCGTGTA CTATGCCGAG TATAACCGAT CGGCGTTATC
167101 GATGAGAGTT TCAAAAAAAT ATAAATCCA CCTAAAGAAA AAGCGATAGA
167151 GAAGGTTCGT ACATGACGCA TCAAGTAGCT GTCTTGCATC AGGATAAAAA
167201 ATTTGATGTT TCGTTAAGAC CTAAAGGGTT AGAAGAATTT TATGGACAGC
167251 ATCATTTAAA AGAACGCCTA GATCTATTTT TTTGCGCAGC ATTGCAACGA
167301 GGAGAAGTTC CAGGACATTG CTTGTTTTTT GGACCCCCAG GCTTAGGGAA

167351 AACCTCACTT GCTCACATCG TTGCCTACAC CGTGGGGAAA GGGCTGGTCT
 167401 TGGCATCAGG GCCTCAGTTA ATCAAACCCT CGGACCTGTT AGGACTTTTA
 167451 ACTAGTTTGC AAGAAGGGGA CGTGT'TTTC ATCGATGAGA TCCATCGTAT
 167501 GGGGAAAGTT GCTGAGGAAT ACCTGTATTC TGCAATGGAA GATTTCAAAG
 167551 TCGATATTAC TATAGATTCA GGACCCGGAG CTCGCTCGGT CCGTGTCGAT
 167601 CTTGCTCCTT TCACTTTAGT GGGGGCAACG ACTCGATCAG GAATGCTAAG
 167651 CGAACCTTTA AGAGCACGCT TTGCTTTTAG TGCGAGACTT TCCTATTACT
 167701 CGGATCAAGA TCTAAAAGAG ATTTTAGTCC GCTCCTCACA TTTACTCGGA
 167751 ATCGAAGCTG ACAGCTCCGC ATTACTAGAA ATTGCTAAGA GATCCCGAGG
 167801 GACGCCACGA CTGGCAAATC ATCTTCTACG TTGGGTCAGA GATTTTGCTC
 167851 AGATCCGAGA AGGAAACTGT ATCAATGGGG ACGTAGCAGA AAAAGCTTTG
 167901 GCTATGCTAT TAATAGATGA TTGGGGATTG AATGAAATTG ATATCAAAC
 167951 TCTCACTACA ATCATCGACT ACTACCAAGG TGGTCCCGTT GGAATTAAAA
 168001 CCTTATCGGT AGCTGTGGGA GAAGATATCA AACTCTTGA AGATGTTTAT
 168051 GAACCGTTTT TAATTTTAAA AGGTTTATC AAAAAGACTC CCAGAGGCAG
 168101 AATGGTAACA CAACTTGCTT ACGACCATT T AAAAAGACAT GCAAAGAACT
 168151 TATTGAGTTT AGGAGAAGGA CAGTGAACT ATTGAAAAAC GTACTTTTAG
 168201 GTCTTTTCTT CAGTATGAGT ATCTCAGGAT TCTCAGAAGT AAAGGTATCC
 168251 GATACTTTTG TGAAGCAGGA TACTGTCGTT GAACCTAAAA TTCGTGTCCT
 168301 TTTATCTAAT GAAAGCACCA CAGCTCTCAT AGAAGCCAAA GGTCTTATC
 168351 GCATTTATGG AGATAATGTC TTATTAGACA CAGCGATTCA AGGCCAGCGT
 168401 TGCGTGGTCC ACGCTCTATA CGAAGGGATC CGTTGGGGAG AATTTTATCC
 168451 CGGACTCCAG TGTTTAAAGA TCGAGCCTGT AGATGACACT GCTTCTCTTT
 168501 TTTTAAACGG GATTCAGTAT CAAGGTTCCC TATACGTTCA TCGTAAAGAC
 168551 AACCATTGCA TCATGGTTTC TAACGAAGTT ACAATCGAAG ATTATCTGAA
 168601 ATCTGTACTT TCTATAAAGT ACCTTGAAGA GCTAGATAAA GAAGCTCTAT
 168651 CTGCTTGCAT CATTCTAGAA AGAACCGCTC TATACGAAAA GCTCCTTGCA

168701 AGAAATCCTC AAAACTTTTG GCATGTTAAA GCTGAAGAAG AAGGGTATGC
 168751 AGGATTTGGT GTGACCAAGC AGTTCTATGG TGTAGAAGAG GCTATAGACT
 168801 GGACAGCTCG TTTAGTTGTG GATAGCCCTC AAGGATTAAT TATAGATGCA
 168851 CAAGGGCTCT TGCAGTCCAA CGTAGATCGT CTTGCTATAG AAGGATTCAA
 168901 TGCACGTCAG ATTCTTGAGA AGTTCTACAA GGATGTGGAT TTTGTAGTTA
 168951 TAGAATCCTG GAATGAAGAA CTGGACGGAG AGATCAGGTA ACCTCTTTCG
 169001 CATGGCTGAT CGCAATTAGC GTGGTATGGG GCTGTAGCGA CACTGTCGGC
 169051 GTTGCTACAT TTTGAGGGAC AAACCCTTGC TGA CTCTCGG CAACTATTTG
 169101 ATAAGGAAGA AAGTTGCTGG AGGCTTTAGG TAAGGTCGCA AGTTGGTCTT
 169151 GAGCTCCCAC GTGAAAAGCA ACATATACAT GCGCTTTTGG CGATTTTATT
 169201 TTAAATGCTA AGAAATTTC AGGGCGCCAT GTCATGGGAT TTCCCATAGC
 169251 ATCTACCCAA CTGATTTCCT TATTGGAAAG AAAGCCTCGA TTAAAAAGTG
 169301 TTTTATATTT TTTTCGAAAC GCAATGAGAT CACAGAGAAA GTGCATCAGT
 169351 GTAGGCTTTG CGGTAAGCTG ATCCCAAAGG AAGTAATTCG CATTCGAATC
 169401 CAAAGCCCAA CGGTTGTTAT TGCCTTCCGC GGTATGGGCA TACTCATCTC
 169451 CTGATTGAAT CATCGGAATG CCTTGCAGAG CCATCAAAGT AAGGAAAAAA
 169501 TTTTCGTA ACT GTCTTTCACG AACTTCAAGA ATGCCAGGGT CTTCTGTTTT
 169551 CCCTTCCGTT CCGAAATTGT AGCTGTAGTT CGCATCTGTG CCGTCACGAT
 169601 TATCCTCTCC GTTAGCCTCA TTATGTTTGT GGTATAAGT CACAGTGTCA
 169651 CATAACGTAA AACCATCATG GCAACTGACA TAGTTAATCG AATTTGTAGG
 169701 CGAGCCGTGA GGATAGATGT CTTGAGATCC TGAAATTCTA GAAGCAAAGG
 169751 TTCCTATGAG ATTTTGATCC CCATTAAGAA ATGCTTTCAC GTTATCACGA
 169801 TACGGGCCGT TCCATTCACT CCATCTTGGA GACAGTGTGG GGAAATAGCC
 169851 CACCTGATAC AAACCGCCAG CATCCCAAGG CTCAGCTATA ATCTTTGTGC
 169901 TCGCAAGTAA AGGATCAAAA GAAATCGCCT CTAAAACAGG AGCGAATTGT
 169951 AGGGGAGATC CCGAAGGACC ACGAGAAAAG ACAGAAGCAA GATCAAATCG
 170001 GAACCCATCG ACATGCATTT CTTCTACCCA ATAACGTAAG ATGTCGAGAA

170051 TCCATTGGGT CGTGGGGGCG CGGTTTGTAT TGAGAGTGTT TCCACAGCCT
 170101 GAATAATTTG TAAAGTGACC TTGTGCATCT AAAATATAAT AGCTCGGAGT
 170151 GTCTATCCAA GGCAAAGAGC AGGTCGTCCC TTGCAAGCCC GTATGATTAA
 170201 AAACAACATC AAGAATGACC TCAATACCTT CTTGATGCAA GGTCTTTACT
 170251 AAAGTTTAA ACTCTCTACT TGGAGCGCAA GGATCAGAGG CATAAGCATA
 170301 ACGTCGGCAA GGAGAAAAGA AATTTAGGGG AGCATAACCC CAATAATTGC
 170351 ACAGATAAGG GAATTTCGAA TTTCTAAAAG GATGCGCAGT CTCATCGAAC
 170401 TCAAAGATAG GTAAGAGTTC AACAGCGTTG ATTCCCAGCT TATGCAGATG
 170451 GTCGATCTTT TCAATGATTC CTAGGAAGGT TCCCGGAGCA TGAACCCTAG
 170501 ATGAAGAAGA TTGCGTGAAG GAACGTACAT GCATCTCATA GATGATCATC
 170551 TCTTCTTTTCG GCAAATGCAG AGGCTGATCA CCATCCCAAG GAAATGGTTC
 170601 TTCTTTTAAA TAACAAAATG CATAATCCCC CTGTTTCTTT CGCGAACCAG
 170651 AACTCTGTGG GGAATGAATA TTCTTCGCAT AGGGATCTGC AAGATATTCT
 170701 TTAAAAGAGT ATTGCATTCC ATGCTTTTTC GGCCCATGAA CACGAAATGC
 170751 ATAAGACGAT TGATCAGAAA TACCCTCGAT CTCTATATGC CAAATCGCAC
 170801 CCGTGCGGTG TGTATCGGGG TAAAGAGGGA CTTCTATGAC TTCTGAATTT
 170851 TCGTCTGTGA AAGCAAGGAT GACTTCGGTA GCTTGTGAAG CATATAAAGC
 170901 AAATCGATAG CGGTTTGGGG AAATTTTAGA AGCCCCAAGA GGTAAAGGAA
 170951 CTGAGGGATA AGAAGAAACT TTTTCCATCG TCGATCAGAT ATCATAACGG
 171001 CGCCAGAAAT ACAAGGGAGT TTTTGATTTT AAGCCTCTGA AATCGTTTAC
 171051 TAACAAATCA CTCTTATGGT AACTTGAAAC AACAACAATC TATTACAAGG
 171101 AGATTCCCTC ATGTCCAGGC AAAATGCTGA GGAAAATCTA AAAAATTTTG
 171151 CTAAAGAATT AAAACTGCCG GATGTAGCTT TTGATCAGAA TAACACGTGC
 171201 ATTTTATTTG TTGATGGTGA GTTCTCTCTT CACCTTACTT ATGAAGAACA
 171251 TTCTGATCGT CTTTATGTCT ACGCTCCTCT GTTAGATGGA CTCCCTGATA
 171301 ATACTCAGAG GAAATTGGCC TTATATGAGA AATTATTAGA AGGATCTATG
 171351 CTTGGCGGTC AGATGGCTGG AGGTGGAGTA GGTGTTGCTA CTAAGGAACA

171401 GCTCATTCTC ATGCACTGCG TTTTGGATAT GAAATATGCT GAAACGAATT
171451 TACTAAAAGC ATTTGCTCAA TTGTTTATTG AACTGTTGT TAAGTGGCGT
171501 ACCGTGTGTG CTGATATTTG TGCTGGTAGA GAACCTTCTG TAGATACAAT
171551 GCCACAAATG CCTCAAGGCG GTGGTGGAAT GCAACCTCCT CCTACAGGTA
171601 TTCGTGCTTA ATTAAGCATA TAAGTACTAC ATCTTTTTTC TAAATTTAAA
171651 AGATTGATAA AGCTCTTCTT AGAGAAGAAA TGACTTTATT GACTTCTCGT
171701 TTTTGTCTCG AATCTGTTTT GAGAAGTTCT TAAGATGTAT TAATTAAAGA
171751 AGTATTAAAA ATAGAAATCT AAAGGCTATC TTATGATGTT TGGGCATTTT
171801 GCTGGTTACC TTGGAGCAGA TCCTGAAGAG CGAATGACTT CCAAAGGAAA
171851 ACGTGTGATC ACTCTGAGAC TGGGAGTGAA GACTCGAGTT GGAATGAAAG
171901 ATGAAACTGT TTGGTGCAAA TGCAATATTT GGCACAATCG CTATGATAAG
171951 ATGCTTCCTT ACTTGAAGAA AGGCTCAGGA GTCATTGTTG CTGGCGATAT
172001 CTCTGTAGAG AGTTACATGA GCAAAGATGG TTCACCGCAA TCTTCTTTAG
172051 TGATTAGTGT AGATTCTTTG AAATTCAGTC CTTTCGGTCG CAATGAAGGC
172101 AGCCGTTCTC CATCTTTAGA AGACAATCAT CAGCAAGTGG GATATGAATC
172151 TGTATCCGTA GGGTTTGAAG GTGAAGCACT GGACGCAGAA GCTATTAAAG
172201 ATAAAGATAT GTATGCTGGT TATGGTCAAG AACAGCAGTA TGTCTGTGAA
172251 GATGTTCTT TTTAATTCCT AGTCATTAAA GGAGAGTTTG TGGTTTTATT
172301 TCATGCTCAA GCCTCTGGGC GTAATCGTGT TAAGGCAGAT GCTATAGTCC
172351 TGCCCTTTTG GCATTTTAAG GATGCAAAAA ATGCAGCTTC TTTTGAAGCC
172401 GAGTTTGAAC CCTCGTATCT CCCCCTTTA GAAACTTTC AAGGAAAAAC
172451 CGGGGAGATT GAACTCCTTT ATAGTAGTCC TAAAGCTAAG GAAAAACGCA
172501 TTGTCCTCTT AGGCTTAGGG AAAAATGAAG AGCTCACCTC TGATGTTGTT
172551 TTCCAAACCT ATGCGACACT AACTCGTGTC TTACGTAAAG CAAAGTGTTT
172601 CACAGTCAAT ATCATCTTAC CTACAATTTT TGAATTGCGG CTTTCTGCCG
172651 AAGAATTCTT AGTGGGGTTG TCCTCAGGAA TTTTGTGATT AACTATGAC
172701 TACCCACGTT ATAATAAGGT AGATCGTAAT CTTGAAACTC CTCTTTCTAA

172751 AGTCACGGTT ATCGGTATCG TTCCCAAAT GCGGATGCT ATCTTTAGGA
172801 AAGAAGCAGC CATTTTCGAA GCGTATATC TCACTCGAGA TCTTGTGAAC
172851 AGGAATGCTG ATGAAATTAC CCCTAAGAAA TTGGCAGAGG TTGCTCTGAA
172901 TCTGGGAAAA GAGTTCCTTA GTATTGATAC TAAGGTCTTG GGAAAAGATG
172951 CCATCGCCAA AGAGAAAATG GGACTCCTAT TGGCTGTTTC CAAGGGTTCT
173001 TGTGTGGATC CACACTTTAT CGTTGTCCGT TATCAAGGAC GTCCTAAGTC
173051 TAAAGATCAC ACCGTCTTGA TAGGGAAAGG GGTCACTTTT GACTCTGGAG
173101 GTTTAGACCT CAAGCCTGGA AAATCCATGC TTACTATGAA AGAAGACATG
173151 GCAGGTGGGG CTACAGTCCT CGGGATTCTC TCGGCGTTAG CAGTTTTAGA
173201 GCTTCCTATA AATGTCACGG GGATCATTC TGCTACAGAG AATGCTATCG
173251 ATGGCGCCTC CTATAAAATG GGAGATGTCT ATGTAGGAAT GTCGGGGCTT
173301 TCTGTTGAGA TTTGTAGTAC CGATGCTGAG GGACGTCTTA TCCTCGCTGA
173351 TGCGATTACA TATGCTTTAA AATATTGTAA ACCGACACGT ATTATAGATT
173401 TTGCAACTCT AACAGGAGCT ATGGTAGTCT CTCTAGGAGA AGAGGTTGCA
173451 GGTTTCTTTT CCAATAACGA TGTTTTAGCT GAAGATCTTT TAGAGGCGTC
173501 AGCCGAAACC TCCGAGCCGT TATGGAGACT TCCTCTAGTT AAGAAGTATG
173551 ATAAACATT GCATTCTGAT ATTGCTGATA TGAAAAATCT AGGCAGTAAC
173601 CGTGCAGGGG CTATTACAGC AGCATTATTC TTGCAGAGAT TTTTGGAAGA
173651 ATCTTCGGTA GCTTGGGCAC ATCTTGATAT TGCAGGTACT GCATATCATG
173701 AAAAAGAAGA AGACCGTTAT CCAAATATG CTTCAGGTTT TGGTGTCGT
173751 TCTATTCTTT ATTACTTAGA AAATAGTCTT TCTAAGTAGT TGCTTTCTAT
173801 TTATTTATGT TTTAGTAATG ACTTTTATTT TAGTTTTTTT AAAATAAAAG
173851 TCATTTTTTT TATTAAAGTT TTCAATCGTC CCTGCCGATA GATCAGGTAA
173901 GTAATTACCT GTCTAATTAG GGAATAAAG ATGATTGGAG CGCAAAAAAA
173951 GCAAAGCGGT AAAAAGACAG CTTCAAGAGC TGTACGGAAG CCTGCTAAAA
174001 AAGTTGCGGC TAAACGTACG GTTAAAAAAG CTACTGTTTC CAAAACCGCT
174051 GTAAAAAAAC CTGCAGTTCG TAAGACGGCT GCTAAAAAGA CAGTAGCAAA

174101 GAAGACTACA GCTAAGAGAA CAGTTCGTAA GACTGTTGCT AAGAAGCCTG
 174151 CAGTTAAGAA AGTTGCTGCT AAACGTGTAG TAAAAAAGAC AGTAGCAAAG
 174201 AAGACTACAG CTAAGAGAGC GGTTCGCAAG ACTGTTGCTA AGAAGCCTGT
 174251 AGCTAGAAAA ACTACAGTGG CTAAAGGTTT TCCTAAGAAA GCTGCAGCCT
 174301 GTGCTTTAGC ATGCCACAAA AACCATAAGC ATACATCTAG TTGTAAACGT
 174351 GTCTGTTCTT CAACAGCTAC GAGAAAGCAT GGCTCTAAAA GCCGTGTTTCG
 174401 TACAGCTCAT GGCTGGCGTC ACCAACTGAT CAAAATGATG TCTCGATAAT
 174451 TTGTGATTTT CGCATTATTG CTCATGTTAA CGGGAAAGGG AAACATTGGG
 174501 TTTCTTCCC GTTTTTCTTT TTAAGGTTAA AAAGCTTTAT AGAGCGAGAT
 174551 CTTCAGGCTT CATGCTGTAC AGTTGGTAGG AAAATACGTA TAGTAGGTTT
 174601 AGGATACTAC TTTTTTGAAT CTACCTATGC AAAAATCCTT AACGAGTTTT
 174651 GATGACTTTT CCCAGGCGTA TGCAGAGAAA GTGCCCGCTA TAGCTCTTAT
 174701 AGGGAGTGCT TTGGAAGACG ATAAAGATGC GCTGATTGAA TTATTAGTCT
 174751 CTGAGAGCTT CAAAGAGCTC GGTGGTCAGG GACTCATGCC AGCAACCCTC
 174801 ATGTCTTGGA CCGAGACGTT TGCACTCTTT CAAGAGCATG AAACCTTGGG
 174851 GATTATTCAT GCAGAGAAAT TCCCTCTAGC AACTAAGGAA TTTCTAAGCC
 174901 GCTATGCTCG GAATCCTCAA CCTCACCTTA CGATTTTGAT CTTCAACCACA
 174951 AAACAAGAAT GCTTTCGAGA ACTGTCAAAA GCCTTGCCAT CGGCTCTTTC
 175001 TTTGAGTTTA TTTGGTGAGT GGCCCGCAGA TCGTCAGAAA AGGATCATA
 175051 GCCTCCTGTT GCAAAGAGCT GAGCGTGTGG GGATTTCTTG CTCTCAATCA
 175101 TTGGCATCTT TGTTTTTGCG TGCACTTGCT TCAACCTCTC TTCCTGATAT
 175151 TCTCAGTGAA TTCGATAAGC TACTGTGCTC TGTTGGCAAG AAAACGTCCT
 175201 TGGATCACTC TGATATTAAA GAGCTCGTTG TCAAAAAGA AAAGGCTTCC
 175251 CTATGGAAAT TTCGAGACTC TCTATTGAAG AGGGATCCGG TAGAAGGTCA
 175301 CCAGCAGTTG CATTTTCTAC TCGAGGATGG TGAAGATCCC TTGGGGATTA
 175351 TTAATTTTCT TCGTACCCAA TGTCTCTATG GTTTACGTAG TATTGAAGAG
 175401 GGATCGAAAG AAAATAAACA CCGAATGTTC GTCCTTTATG GAAAGGAGAG

175451 ACTACACCAA GCCTTAAATT CTTTATTTTA TGCAGAAACC TTAATTAAAA
175501 ATAATGTCCA AGATCCTATA GTAGCTGTGG AAACTTTAGT GATTAGAATG
175551 GTTAACCTGT GACTTTATAT CTTCTTCCCA ATACTCTCGG TACCCGTGCT
175601 GTAGAGACTC TCCCCTCCGT TATAGGAGAA TTAGTTCATA GACTAGATGG
175651 GCTGATTGTA GAAAGTGATC GTGGGGGTAG GGCATTTCTA AGTTTATGGA
175701 AAATTCCCGA AGTTCATAAA TTTCCTCTTG CTATTCTTAG TAAACATGCG
175751 CGCCTCCCTA AGGCTTGGA TTTTATCTA GAGCCTATCG TAAAACACGG
175801 GGAGAAATTGG GGAAGTATCT CTGATGCGGG TCTTCCCTGT ATTGCAGATC
175851 CTGGAGCGAG TTTAGTGCGT CGTGACCGTG CTTTGGGGAT TCCTGTGCAG
175901 GCTTTTTCAG GTCCCTGTTC GATAACGTTA GCGCTCATGC TTTCAGGCTT
175951 GCCTTCCCAG AGCTTTACGT TTTTGGGATA CCTCCCGCAA AGTCCTAAGG
176001 AACGTGTAAA GTCGATAAAA AAGGCAGCGA CCTCCAAAGA GGTATCTACT
176051 TCAGTATGTA TAGAACTTC TTATCGTAAC GTCTATACTT TTGAGTCTCT
176101 TCTAGATACT TTACCTTCCT ATGCGGAGCT TTGTGTTGCC TCTGACCTTT
176151 CGGGCCCGAG TGAAGTTGTT CTCACACGCC AAGTGCAATC ATGGAGAACT
176201 ACTGAGGACT TAGGTTCTGT GAAGCAAAGT ATAACCAAAG TCCCAACAAT
176251 ATTTCTTTTC CACATCCCGA ATTAATTTTT TCTTAGCTTG TTTTATCCCC
176301 GGTATCATTG TCAGGAGTTT CTTAGCATGT AAGCTCTTCG GGGATATTTT
176351 TCTTGTGATC ATAAGCAAAA ATTAAAGGTT GCATGATTCT AAATAGCGCA
176401 TAATCATAGG TTTGCGTCTT TTTAAAATAA TCGTAATCCT CTATTGCGTG
176451 TTATCAATAC TTATTTTATA ATCTTTGCCA AATAAGTTTA GAGCGTTGAT
176501 GACAGCACCC ACAGAATCTC GATCTTCACC ACCAACACTA CTCGAAGAAA
176551 CGGAACCTCT ATCCCCAAAT CCTATTCCCG CCGATATCCA AATCCCGAGA
176601 ATTACAATAT CTCCTCCTTC TCTAGACGTA TCAACAGTTG CATCTTCGGC
176651 TGAAGATATC TCGGTCTTCA TTGCAGGAGG CCCTAGAAGT TCTTCATCTG
176701 CTTCTGTGGC TTCGGATGTG TATGAGTTGG TTTGCCTTTG TGGCGGTGAT
176751 GAGGATCCTG AGCCCCCAGA TTCTGAGGTA CGTACTCTCT ATGTAAATGG

176801 AAGTTGGCAA ACGCATCAAG AAGCTGTACA GGAATTGCTG TATATTAGTG
176851 AAGTACGGGG AGAGGCCGTT CGTCTTCTCT ATAACGATGG AAGCGGCATG
176901 TCTCCTTGGC CCATCAGTCC TTGTCGTA CTTCCTACTC TCGATCATCC
176951 TTTATGTCAG GCCCTTTTGA CGGTTTGGGA ACAGTTTTTC TCTGCTCCTG
177001 AAAATCAGAA TCGTGAGTTT CTAGTGATTT TCTATGGGGA TGCATCGCCT
177051 TATATACAAC AGGCGTTAAC GCAATCTAGG CATAGTCCAC GTATTGTTGT
177101 TGTAGGGATT TCCCCGACGG TCTTTATTCA AGGAGACTTT AGGGTCCATA
177151 ATTACCGTGT TTCTGGAGAC TTCTTTAGTT CTCTGGATTG TCGGGGAAC T
177201 AGGGCGGAGA ACACCACGAT ACTGCCGTAT TCTTCGGGTC TTGAGGGTGT
177251 TTTTCTGCCT TCTATCCGTT GTCCTTCTTT TACTTGGGCG GTGCGTTTTG
177301 GAGAGCAATG CTTGGTTGCG AATAGGGGTG AGGATGTAGA AGATAGGGGA
177351 GGTCTTTCTC AAGATGCCGA AAGATCACAG TTACCACACA GTGAAAGAGA
177401 TCTAGCTGTT GTCATTGATT CTACGGATCC TAGTTCTATG AGTAGGCTTG
177451 TAGAATGGTT GAATCAAGGA TCGCCTTCAT CAGATATGGA AATCAATCCC
177501 TATCCCCAAC GGTGTCCTGA TG TAGCTCTT TCTGCGCTTT ATGCAATTTC
177551 TAGAGTTTCA GGA CTGCGC AGGAATGGAT CCTAGCCTCT GTTCATGAGG
177601 GCTTAGACTT GCAGATCTGT TACTCTTTAA TTTTGATGCA CACGACGTTT
177651 GCGGTCCGGT ATTTTTTCTT ACTCTTTACA AATTATCCTC AGTCTAGAGA
177701 GAGATTCCGT ACTGCACGAA TCGTAGCACA ATCTCTATAT TTACCAAGCA
177751 TCCTTGTTCT TGTTTTTGAT TGTGGCAACG TCCTGCGTAA ACTATGGATG
177801 CCTCAGGAAA TCTTACGAGC AATTTTTATT TCTGCGTCTA CAATTCAGG
177851 GAGTATTGTC TTTGTAGAGT GCACTCGCTG GATGGGGCGA GGTCTTAGAC
177901 ATCGTG TACA ACAATTTGTG CAGCAACGAG TTATAGGAAG TGGCCTGCCT
177951 GTAGGAACAG TACGAGCTTC TTATCGCGAT CGTGCAGGCT TTATCATAGG
178001 CTTTTTACAA ACTGTACATG GAGGACTTTA TTTGCCGGTA TCCATTATGG
178051 TGCTTAACCA GATTGCAATA CAAGTTCCAC GTATCTTAGT ACGTCCAAAT
178101 AACACTGCTG TTTATGATCT ACATAATAAA AGTGCTGAAG AAAATTGGAG

178151 CAGTGGTGAT GTATTAGCTG TTGGCCAAAC ATTAACTTC ATCTTATGTG
 178201 CTTTCGTCTT GTTCGTAAAT CTATGGTTCT TTGTGAAGTC CGTATTACGC
 178251 CATTCTAGGC GTCGTCGTCG CTAATGTATC TTTGGAGACA TCTTGGTATT
 178301 TTAAACATAA GAGCACGGCA AAAAAAGAA AGACCGAGAA AGAAGCTTTA
 178351 CCTTTAAGAG TTCTCTGATG TTTCCCGGAG CATCAGGAGT CCGAAGGGCA
 178401 GAGTCTAGGT TTTAGCCTTC CGCAGGGATT AGAAGGAATA TCCCATAATC
 178451 TCTTCCGCTA TTGTATCGTG GAAGAGACGG CCCTGTCTAT TTAGGGCAAG
 178501 ACATTGTCCA TGCACACTGA ATAGGTTTGT TAATTTTACA TCTTGCGTAA
 178551 GCATGGAGAT AAGTGTGGAG GGGAACTCCG CGAGGTCTGC TCCTTCAAGG
 178601 AGTCGGAGTC GCAGGGCTAA GGCTTCTTTG ATTCGTTCTT TTTTGGGAG
 178651 AATTTCTGAG GTCTCTGGG TAGGGAGATT CTTACGTACA GCACGTAGAT
 178701 AGTGAGAAAT ATGACTATAA TTTTTTGACC GCTCTCCGTG AAGGTATTGC
 178751 GAAGCTGAAA CTCCTAAGCC TAAGAAAGGG CGATCTGTCC AGTAATAGAG
 178801 GTTGTGCTTT GCGGGTAAT CTGGCTTGGC ATATGAAGCA AGTTCATAGC
 178851 GTTGAACCC TTGGGAGAGT AGGAGATTTT CAGCAAGGAG GCTCATCTCA
 178901 GCTAGAATTT CTTCTTGGGC AATTGTGGGG ACTAGAATTT TGCGGTGTTT
 178951 ATAGAAGGAG GTGTGGGGAT CTATAGTGAG GTTGTATAGA GAAATGTGAG
 179001 TGATAGGGAG AGTCAGAGCT TGATGTAGGT CGCTTAGGAA TATCTCCAAA
 179051 GACTGTGTGG GCAGTCCGTA GATTAGGTCT ATAGAAAGAT TAGAGAATCC
 179101 GTGATTCTGG CATTCTTGCA GTGCTGTGAT TGCCGCAGAT GAAGAATGCG
 179151 TTCTTCCGAG GAGCTGTAGG ATAGAGTCGT CGAAGGTTG TACGCCAACG
 179201 CTAATTCTAT TTATTGGAGT CTCTTGTAGT TGACGTAGAT AGCTTACGGT
 179251 GAGATTTTCG GGGTTGGCCT CTAAAGTAAT TTCCCGGGCA TGGGGGGCTA
 179301 GCTCTTTGAG GATGCGCTTA AGATCAAGAG GAGAACTAA TGAAGGTGTT
 179351 CCCCCTCCAA AAAACACAGT CTCTATGAAA TGCGTCTCTT GGATGGGGGC
 179401 TAGCTTTCTT AGCCCCTCTT GAATTACAGC ATTACAATAG AGCGATACAG
 179451 ATTCACTTTT GTAGGGGATT GTATAAAAAC TGCAATAGCG ACATTTTTTTT

179501 GTGCAGAAGG GAATATGAAT ATAAAGAGCT AGGGGAGCCT TACCATTTCAT
179551 TAGCATCGGG ATCAATGATA CCACCGCGTC TCCAGCGATT GCGATCGCTG
179601 AAAGGGTCTT CATCGTCTTC GTGGGAATAG TCGACTTCTA CTGCCCCATC
179651 TTCACGTCCT TCATCGGTGA GTTCCAGTTC TACGGTTTCG CCTGGGTCGA
179701 TATCAATTTT TGTTCCTGTG GGTCCACAA ATTCAATATC AGACATACTA
179751 GGACCATCAG GGTAGACAGT CTTGCGGGA CTGCGTCGTG GTGTGACGTA
179801 CTCTTCAATC TCTCCGTTTT CTTGAGGAG TTGTAGACGT TCTTTTTCTT
179851 CGTAAATTTT ATGTTCTAAG ATTCGGATTT CTTCTTGGTG CCTGCTAATT
179901 TCTTTTTTTG GCACTAAACC CAACTGCATC CATTGCGTAA GGTCATGAAG
179951 TTCTGATTCT AATTTTTTAA GACGTTCGCT TTTCATCCAA GGGTTCCTAT
180001 CCTGTATTCG AGAGCAGGTG CATAGCATTT TTTTGTTTTT TCAACTGCAG
180051 AAAAAAACA AAGATTTTTT TTTTTTTGAA AAATAAAAAT TTGCTATGAA
180101 GCAGACATTT AGATCGTATT TATTGAGTTT AATTATTTTA TGGATTCCGA
180151 GTTTGTGGGG CAAGTATATT CTTGCGATAT GGATTGGATC GAGTCTATGT
180201 ATCAGAGATT TATGAATCAC GAGACTTTGG ATCCTTCTTG GAAGTATTTT
180251 TTTGAAGGGT ATCAGCTCGG TCAAGCAGCA TCTCCATCAG AAGCTAGTAC
180301 TAAGATTCTT GGGAATGAAA CTATTGCTAT GCTTCAAGAA CAAAAATCTC
180351 AGTTTCTATG TACGATTTAT CGTTATTATG GATATTGCA AAGTCAAATT
180401 TCAACGCTTG CCCCAACTAC AGATTCTCGA TTCATTCAGG AAAAGATCGC
180451 TAAGATTGAT CTGGATGAGC AGGTGCCTTC TGCGGGTCTA CTCCTAAAG
180501 CTCAGGTTTC GGTACGAGAG CTGATCGAAG CTTTAAAAAA ATGCTATTGC
180551 GGAAGTCTTA CTTTAGAAAC CCTAACATGT ACTCCTGAGT TGCAGGAGTT
180601 TGTTTGGAAT CTTATGGAGA AGCGACAAGT GGAGCGCTTT GCAGAGCAGC
180651 TCCTTCGCTC CTATAAAGAC TTATGTAAAG CAACGTTTTT TGAAGAGTTC
180701 TTACAGATAA AATTACAGG TCAGAAACGT TTTTCTTTAG AGGGCGGAGA
180751 GACCTTGGTC CCCATGTTGG AGCATCTTGT TCATTATGGA TCGGCATTAG
180801 GAATTTCTAA CTACGTTTTA GGAATGGCCC ATCGAGGTCG TTTGAATGTA

180851 TTAACGAATG TTTTGGGAAA GCCTTACCGT TATGTCTTTA TGGAGTTTGA
180901 AGACGATCCT GCAGCACGTG GTTTAGAGAG TGTTGGGGAT GTAAAGTACC
180951 ATAAAGGGTA TGTGCTAAAG TCCCATCAGA AAGATAGGGA AACTACCTTT
181001 GTGATGTTGC CAAACGCTAG TCATCTCGAA TCTGTAGATC CTATTGTCGA
181051 GGGGGTCGTG GCTGCCTTGC AACACCAAGG TCACGCAGGT AAAGAGCAAA
181101 GCAGCTTAGC AATTTTAGTT CATGGAGATG CAGCATTTTC TGGTCAGGGA
181151 GTGGTTTATG AAACCTCTCCA GCTGAGTCGT GTTCCAGGGT ATTCTACTGA
181201 GGGTACGCTT CACATTGTTG TGAATAATTA CATAGGGTTT ACCGCAGTGC
181251 CACGGGAGTC AAGGTCCACC CCTTATTGTA CGGATATTGC TAAAATGCTA
181301 GGGATTCTCTG TATTTTCGAGT GAATAGCGAG GACGTCGTTG CCTGTATAGA
181351 AGCTATAGAG TACGCTCTGC AAGTTCGTGA GAGATTTAGT TGTGATGTGA
181401 TCATAGATCT CTGCTGTTAT CGCAAGTATG GACATAATGA AAGTGACGAT
181451 CCCTCAGTAA CAGCTCCCTT ACTCTATGAT CAGATTAAGA GAAAGAAGAG
181501 TATTCGCGAG CTGTTTAGGC AATATCTGTT GGAAGGGCAG TTTGCAGATA
181551 TTTCTGAAGA AACTTTGGCA TCTATTGAAA AAGAGATTCA AGAGAGTCTG
181601 AATCGTGAGT TTCAAGTATT GAAAGGGACG GATCCAGAAC CCTTTCCTAA
181651 AAAAGAATGT CATCACTGCG ATCGCTTAAA TAACGGCGAG CTTATTTTGC
181701 ATGATTGTGA TGTTTCTTTG GATCGCGAGA CTCTTTTTC AATGAGCTCG
181751 CGTCTTTGTG GTTTCCTTGA CAATTTTCAT CCCCATCCTA AAATTAAGAC
181801 TCTTTTAGAA AAAAGAATGA AAATGGCAGA AGGTGGGGTT GGTATGATT
181851 GGGCGATGGC CGAAGAATTA GCCTTTGCTT CGCTATTAAT CGAAGGGTAC
181901 AACCTGAGAC TCTCAGGTCA AGATTCTATT CGCGGGACAT TCAGCCAACG
181951 ACATTTGGTA TGGAGTGATA CTGTGACTGG AGATACCTAC TCTCCATTGT
182001 ACCATCTTTC TGCAGAGCAG GGCTCTGTAG AAATGTATAA TTCTCCTCTT
182051 TCCGAATATG CAATTTTAGG GTTTGAGTAT GGCTATGCTC AACAGGCATT
182101 AAAGACTTTA GTGTTATGGG AAGCGCAGTT TGGGGATTTT GCTAATGGTG
182151 CACAAATCAT TTTCGATCAG TATATCTCTT CGGAATTCA GAAGTGGGAT

182201 TTACACTCTG ACATTGTTCT GCTTCTTCCC CATGGGTATG AGGGCCAAGG
182251 ACCCGAGCAT TCTTCATCTC GTATAGAACG TTATTTGCAA TTAGCCGCGA
182301 ACTGGAATTT TCAAGTGGTC TTGCCTTCCA CTCCTGTGCA ATATTTTCGG
182351 ATTCTCAGAG AGCATGCTAA GAGAGATCTT TCTTTGCCTT TGGTGATCTT
182401 TACTCCTAAG TTGCTGCTGA GATATCCACA ATGTGTAAGT AGTATCGAGG
182451 AGTTCACAGA ACCTGGGGGA TTCCGTGCTA TTCTCGAAGA TGCCGATCCT
182501 AATTATGATG CTTCTATTTT GGTATTGTGT TCGGGAAAAGA TCTATTATGA
182551 TTATGCAGAA ATGCTTCCTC AAGATCGGCG TAAGGACTTT TCTTGCTTGC
182601 GTATAGAGAG CTTGTATCCT TTAGCTCTTG AGGATTTAGT GAGCCTTATC
182651 GATAAGTATT CTCATTTGAA ACATTTTGTT TGGCTACAAG AAGAATCCAA
182701 GAATATGGGG GCCTATGACT ATATGTTTAT GGCCTTGCAA GACATTCTTC
182751 CTGAGAAACT GCTATATATA GGACGTCCTC GGAGTAGTTC CACAGCTTCT
182801 GGATCAGCGA AGCTCAGTCG TCAAGAGCTG GTCACGTGTA TGGAAACCCT
182851 CTTTTCTTTA AGGTAAATTA TGA CTACAGA AGTACGCATT CCTAATATTG
182901 CAGAGTCGAT TAGCGAGGTG ACCGTAGCTT CCTTGTTAGT TACAGAGGGT
182951 GCTCTGATTC AAGAAAACCA GGGCTTACTA GAAATTGAAA GTGATAAGGT
183001 AAATCAGCTC ATTTATGCCC CAGTATCGGG AAGAATTTTC TGGGAGGTTT
183051 CAGAAGGCGA TGTTGTTCCCT GTAGGGGGGG TAGTGGGAAA AATAGAGCCC
183101 GCAGGTGAAG GGGAAGAGCT TGGAGATTCT CAGTCTAAAG AGACTATAGA
183151 AGCTGAGATC ATTTGCTTTC CTCAGTCTGG GGTGCGTCAG TCTCCTCCAG
183201 AGAATAAAAC GTTTATTCCT CTTCGTGATC AGATGGACCA AGGATCCCAA
183251 GGTCTTTCTG CAGGAGATCG AGGAGAAACT CGAGAACGCA TGACCTCGAT
183301 TCGTAAGACA ATTTGCGGGC GTCTTTTGTC TGCTTTACAT GAGTCTGCGA
183351 TGCTCACGAC ATTCAATGAG GTCTATATGA CACCTCTTTT TCATTTGCGA
183401 AAGGAAAAAC AAGAAGAGTT TCTATCTCGA TATGGGGTGA AGTTAGGATT
183451 TATGTCTTTC TTTGTGAAAG CTGTCTTAGA GGCTTTGAAG GCATATCCAC
183501 GAGTGAACGC CTATATTGAT GGCGAGGAGA TTGTTTACCG TCACTATTAT

183551 GACATTTCTA TTGCTGTAGG TATCGATCGA GGACTTGTGG TTCCTGTGAT
183601 ACGCGATTGC GATAAACTTT CTAACGGGGA GATTGAGCAG AAACCTCGCAG
183651 ATCTTGCCCT TCGGGCTCGT GAAGGCCTAC TTGCAATAGC GGAGCTTGAG
183701 GGAGGAGGTT TCACAATTAC CAATGGAGGC GTATATGGAT CGCTACTTTC
183751 GACTCCCATT ATCAATCCCC CGCAAGTGGG GATTTTGGGG ATGCATAAGA
183801 TAGAAAAGCG CCCCGTTGTT CTTGATAATG AAATTGTAAT TGCAGATATG
183851 ATGTATGTCG CTTTAAGCTA TGATCATCGT CTTATTGATG GGAAAGAGGC
183901 TGTGGGTTT TTAGTCAAAG TGAAAGAAGG CCTAGAGAAT CCTGCCTCAT
183951 TACTCGACTT GTAATTTCTC TGATTCTCAT AAAGGCTCTT TTAGAGCCTT
184001 TTCAGATTTT TTAACCTCTT TTCTTATCAT GAAAAGGATT GCACTCATCT
184051 TGAGTAAGGA ACAATGTAAA TTGAGTATCC GCACTTACAT TGGCTTGATC
184101 CGCAGCCTCA TTCAAGGCGC CTAACAACTT TCCGGGTAAG TACACATTGT
184151 CTCTTTTAAA GTGAGGTTTC CCGGTATGAT CAAGAGGAGT GAAAGCTTCT
184201 TCTGCGTGAA TATGTACTGA TGTGAAGCCA TTTTCTAAGA GATCTACAAA
184251 CAGATTTTTC AAGAACTCCT TTGAGTCTTT TTCAGAAAGC GAAGAGCTCA
184301 TCGGATCTTT TCTTTCTAAG ATATCTTGTA CATCCTTGTC GTAATAACCC
184351 CAAGGAGTCC ATGCTAAAAC TAAAGTTGAA CTGACACCAG GAGTTGGAGA
184401 ATTGAAAATT GCCCATTGAC CTGCTTTTTT TCTAGAGTTC ATTTTAGTTT
184451 TCATAGCTTC CCACTGTTCC TCTCCTAAGA GATCTTGACG TTGCTTTAGG
184501 GGAGCCGAGA CTTCTTATC CCCTTTCCAA TGCGAAGTTT CTGGTGAAAG
184551 GAGGAATAGC GAAGGTACGC GAGCCCCGTG ATCTCTAATT TCTGTGAGAT
184601 TAGGAGCTTT AGCTGTGACA ACCTTAAATT AAGGACCCTT CTTTAGATGG
184651 TGAATTTAC AATTCGTAGC TTCTGTAATT TTGATTTCTT TAGTGAGTTT
184701 TTCTTTGGGA GGAAGAGCGT CGTAAGGTAA GGAACAATGA TGTTAATAGG
184751 GCTAGTTTAG TTAAATAAGA ATAGGGGATC TCAGGTTTAG GAGTGATGGG
184801 TTGTTGCTCG ATTACAATGA TGGTTAATGG AAGAAGAGCG TTTCTAGAAA
184851 GTGCATCTGC TTCCTCGTCT TCTTTAGGCA ACAAGCTCAA GATGAGAAGC

184901 ACGAGCCCGA TGGCTAATCC CACGAGTGTG ATAATACCAG CAATGCTATA
184951 CGCCTGTATT AGAGGGAAGC CTAAAGCTAA GAAGGTAACG AACGTGATTC
185001 CAGCTGCTAA CATTAGTGCA GCGAAAAGTA TGTAAGAAGC AGCACGCGCA
185051 ATTTGAAGAG CAAGACTTTT ACTTCTGAGT GTGCCTTCTG TAGCTTGGGT
185101 TTTGTTAGCT TCAATCTCAG CGAGTTGCGC TTTTTTTGGC TTCACTGGTT
185151 CGTGAACCTT ATGATTAGTT GATGCGCCTA ATTCATAGA AAGTCTTTAA
185201 GCTTAATTTT TTCCTAAAAA TTGTAGTTTA CATGTTGTTT TTCATCAAGA
185251 AGATTACTGA ATTCTGAGTG GGTGGAACG AAGCATGTTT TTCTAGAAGC
185301 TAAGGTCCCT AGGGGAAAGG AAAAGAAGCT TGGCTGGATC TTTTAAATCT
185351 CTGGGTAGGA GAAGGACGGC GGTACATTA TTTCTATTTT TTGAATGTTG
185401 GCCTGTGAGA TTGGGATCGG GATGGTGAGC AAGGTACCTT GAAGAAAAGA
185451 AAGCGTCGTG TTCCGTGTAG GTACAGAGGT CTGAGATAAA AATGCGGTCT
185501 TTAGAGATTC CTAATTCGT AAGTTGCTTG CGAGCAATCG CACGCAGGTC
185551 AAAATGGTTT TTGGGATTCA TAAAGGGAAG AAAGCTACGA GGAAATAACG
185601 TAGCGTAATC GGGATAGATA GCATAATCTG GACCGATGGA AGGGCCGATA
185651 GCTACGAAGA GATCTTGTGG TTTTGTATGA AATAATTTTT TCATAGTACC
185701 TACGGTGACA GCATAGATAT TGCCAAGCAA TCCTCGCCAT CCGCTGTGTA
185751 CATTTGCGAT TGCCTGGTGT TCTCGATCAT AAAAGATAGC TGCTTGCAA
185801 TCGGAATGGC GGATATGGAG AGAGAGGAGC GGAGACTGCG TGCACAGTCC
185851 GTCTGCAGGT TGGTAGGTGG GGGATGTAGG TGTAACACAA CGTACGGAAG
185901 TGCCGTGGCG TTGATGAAGG TCGCAATACT TCGGAGATTG GAGAGCTGAA
185951 GCAATCTCAG GATTCTTGGC TGCGAAGACC GTGCCCTCGG CATCCTTTTG
186001 TTTTGAAAAG ACTCCGTGGC GTATAGGCAA ACCATCGAAT TCTTCGAAAG
186051 TCCAATAGTT CAGAGGGTGA GTGTGGAAGG ATAGAGTCAT ACTGTCAACT
186101 TAGTTTCTTC AGGGTCTTTC CATACCCCAT GTTCTTGTA AAGTCGAATT
186151 AATTCTTCTT CAGCATCTTC CATGGGTATG TGAGCTTTTA CACAAGTATG
186201 TTTTACATAA AGATCGATCA TCCCTGTTTT GGAACCTACA AATCCAAAAT

186251 CTGCATCTGC CATTCTCTCCA GGGCCATTCA CAATACAACC CATGATAGCG
 186301 ATCTTAAGTC CTGGTAGGTG CTGCGTTCTC TTGCGGATAC GTGTGGTGAC
 186351 TTCTTCAAGA TCAAAGAGGG TCCGACCACA CATAGGACAG GAGATGTACT
 186401 CTGTTTTTAC AAGGCGCACC CCTGCATTTT GTAGAGTGCC AAAGGCAATT
 186451 TTTAGCACGT CCTGTAGGGG AAGGTTCGGT AAGTCAAGAA CCACAGCTTC
 186501 TCCAAGGCCA TCAAGAAGCA GAGCTCCAAA CTCTGTTGCT ATGGAAATAG
 186551 CAGCTTCTTC TTTATTGTCA AAGTCCCTTG AAAATACTAG CTTGGTCGGT
 186601 TTTCTTGGT GTCCTTGTTT TTCAAAGAAA TCTCGGGAGG TATGAATGAA
 186651 GGGGTCTGAA GCATGAAAAT GCACAAATGG AGCTTGATGA ACAGCAGGGC
 186701 TATCCCAAAT CTCCTCATTG TGTTTATATA GGCAAGGCAC TTGATGGTGG
 186751 TGGAAACTA AAAAGTGTTC TCGAAGTACA TCTGTAATAG GAGCATCTTT
 186801 TAACTCAGGG GGAACGACGA CCCCTTCAGG AGTTGTGAAT GCTTTTTCTT
 186851 TTGTTACGGG ATTTACCCCC AAGTGTCTA AGAGTTCTTC AGGAGTAAAG
 186901 TCGGTAAGAT GGTGAGGATA GAGTTTTAAA AAGACTCCGT AGACGTCTCC
 186951 CCAAAGTGTT GTTTTCGCAG GCTTCTCTGC AGCAGAAACA AAGTTTTCGG
 187001 AGTGTGTAG GGAAAAGGA TTTTCTTTT CTGGAAGGTC TAAGTAGATT
 187051 TTCGTATGGC GTAGCAAGCT ATCACAGACA GGAATTTCTG TAGTGGGACA
 187101 CCCTGTGAGA GAGCAGCGTA TGGTATCCCC GAGTCCTTCG GCAAGAAGAG
 187151 TTCCGATTCC TACTGCGGAT TTTATGATCC CGTCCACGCC CATTCCAGCT
 187201 TCAGTAACTC CAAGGTGAAG GGGATAGAGC CAGCCTCTAG CATCTAAGTC
 187251 TTTAGCAAGT TGGCGGTATG CAGTTACCAT GATCTTCGGA TTGCTAGATT
 187301 TCATTGAGAA GACAACATCT CTATAATTCA GCTTTTCACA TACAGCGATA
 187351 TATTCAATTG CTGAGGCTAC CATTCTTCG ATAGTGTCGC CATATTTTGT
 187401 CATGATTCTT TCGGAAAGTG ACCCGTGGTT CACTCCAATG CGCATAGCCT
 187451 TGCCTAGTCG CTTACATTTT TCTACTAAAG GAGCAAACCTT TTCTTCAAGA
 187501 CGCAGGAGAC TTTGGGCATA GCTTGCCTCT GTATAGATCT TCGTCCCCTT
 187551 GAACATGTTC CTCTTATCTA TGTAAGTGCC TGGATTGATG CGAACCTTGT

187601 CAGCAAAATC AGCAACTAAC ATAGCTGCTT GAGGGAAGAA GTGGATATCT
 187651 GCAACCAAAG GGATATTTAA CCCTAGAGCA ATCAGACGTT CTTTAATTTT
 187701 TTCACAGGCT TGTGCTTCCT TGATTCCCTG TACAGTCACT CTGACAATAT
 187751 CACAATTATG TTCCGCTAGA GCGTAGATTT GCTCTACTGT ACTGTCAATG
 187801 TCTGTGGTTA ATGTCGTTGT CATTGATTGG GTTTTTATTG AGTGGTCACT
 187851 GCCTATGTAT AAGTTGCCTA TTCTTACTGT ATGGGTTTTG CGTCGCGAGG
 187901 AATTGATGGC AGGGGTAATG AGTGTCAATA AAATCTCAA AATTTCAAG
 187951 TTTTATCAAT TATAAAAGCC GAGAGAATTT TGTGAAGCT GATAAAGCTT
 188001 CTCTAAGGAT GCCTACATTT ATCAATTTTA CAGAACTTCA CCTCAATAAG
 188051 AGGAATTTTC AGAGAGAAAC GACGACCTTG GATGGCCGAG TGTTTCAGGC
 188101 CATCAAGAGA GTATATGGAG AGATTATTGT GAATTAAAAT AATTATTTCA
 188151 TACTTATCAA GGGGATAAAG ATTTTGCTTT AGAGGACGAT TCGTAAGGAA
 188201 CTCGAATCTG TTTTGTGTGTT CGTAGAACGA ATCCCGAAGT TTGACAAAGA
 188251 GAAGGAGTTC CCAAGAGGAG TGCTTGGGGC TCGTACTTTG GGAACGCTTT
 188301 TTTCTAATTT GAAATCGGTG TATTTTGTGA TGGATGCGTG ATTTTATACA
 188351 TCGAGATTAA CATGATAGCC AGAAGGATTA GGGAGAGCAA GAATAAGGGC
 188401 CCTATAACTG TAAAGTAGGC TCCGACAGAT GTTGCTACAA AGATCATGGC
 188451 TATAGCAACG ATTCCAGAAA CAAAGGTATT GAGGAGGAGT AAAGCAGTTG
 188501 AAGCCGCAA GGCAATTTTT ATGCATCGTC CTGATCCCGA TAACCTTTCT
 188551 AGGAATTCTC CAAGCTTTGT TTGTTCAACT GGAGATGCAC TTGATGTTCC
 188601 TGTGACTACT GGAGAAGACA TGTGATTTCC TATATTTCTA TGACAAGCCC
 188651 TATAATTTTA TAGTTTTCTA TTTTAAATGC AATGTAATTA GTGTTTTACA
 188701 AATTAATCTC CTACTAAAAT AGGAGGAGAC ACTCTCATAT TTTCTATGAC
 188751 CTAGGGAATA AAATATTTTA TTTATAATGA GTGAAAACA CCGCTGTTTG
 188801 TCTAAGAAAT ACTTTTAGAA GAGAGTTTAC GTTTTATCTG CTGTTCAAGCA
 188851 GTTGTTCCCT CGGTTTTGTT TACAAGGCA AGCTTTATAA CAAGAAGAAG
 188901 CACACAGAGT ACTGTTGTAA GAATAAGGCC TCCTAGGATA ATCAAGCAGA

188951 TGGGAGGAGT CCCAGGAATA AAGAAAGCTG CAATTATTCC TAAAACTTCT
 189001 ATTACTAGAA TGGCAACAAG GATTTTAACT AAAAGTTTAA TCGTCTTCGT
 189051 CGTTTTATCT TGATGCTTTG CTTTGTAGAGC TTTGGAATTT GGTTCATGA
 189101 GGCCTAGATT TCCAGAGGAA GGGCGGTGGC TTGTGGGTAG GGGGGCTGAG
 189151 GACACGGGCA TCGTCTTTTA GTTAGAATTT TCTTTATTTT CAAAAGCTT
 189201 TGTCTTATGG AGAGCAGAGG CGTTTCTATT TGCTTGAGAA AGTATAGCGT
 189251 GGTTTAGAAA GTTAGAAAAG ATTTTTTAAA AAATCTTCTA TGTTTTGGAT
 189301 GTTTTAAAGA GAGAAGTCTT ACCATTTTTT AAGAAAGGAT TTCCTTTTCA
 189351 CAAAAGGCA TCTACTCTCT TTTGCTTAAC CAGGGGGAGA GATTAGGGTG
 189401 TGATGAGGGG ATGTCTAGCA AGAGTTGTCA AAATGATAAC CCACGGTTGA
 189451 TTTTGATTTT CTGCTTCTGA TCCAAATGTC TGCAGAGCAT CTATCAAGGC
 189501 TAACTTCACA CCATCAATCC ATTGCAGACG TAGACTAGTA GTCTCTTCGT
 189551 CTTTTGGAGA GCCAGGGAAA TGGGAGGAAA GCAAGGGGAG CTGAACTACG
 189601 GTTGTTTTAC GGCCTTGGC CTCATTCAA CAGTTAAGGT AGGCTTGCCT
 189651 ACAAAGGTA AACGCATCAT TAGGATTGTA GTTATAGTCA GAAGCTTTCG
 189701 GCCCTAGAAG TTGTCCTAAG AATTGCGGTA GTCCTGCTTT ACCTGCATTC
 189751 GTGGTTCCGT TTAGATTTTC CCATTTTGCT GAGCGGCATT CTCCTAATTG
 189801 TAGGGGCTGT TTTGAACGGA GAGGATCCGG AGATTTTTTC GAATTTTCCC
 189851 AACAGCGTAC ACTAGTTGCT TTCGCAAGTG CGAGACTTGT TCCTTTTACG
 189901 TTGTTTGCCA TGTTTGGGGT GGCTGCATTC ACAATCATGA CGATGCCTTG
 189951 AGTTTTATAG CGAGGCTTTT CAATGGGCCC TAACGTAGAG ATCAAAGTCA
 190001 GGTTTGAGTT TTTCAAATGC CAAGCATAAC CGGTTTGATT GTCAGCAGCG
 190051 ATGAAAGGCA TATCTGTGTT AGCCTGTAGA TCAGGTATAC GGTCCTAATT
 190101 TTCTTGTAAC AGTGTTTGAT GGGTTCTAGG GGAAGGAAGC GGGGGCTCTT
 190151 CTAGGACAGC TTCCGTAGGC ACTGGGGTGA GTGCTAGGGG CGGGACAGGA
 190201 GGAAGGTCTG TATCTTTTAT GATGGGCGTC GGCTGATGCG TTACGCTTAT
 190251 AGGACTTTTA GGTCTCTTA AGAGAAAATA GAGAAGAATA CTAAAAGCCA

190301 AGACTGCTAG AGCGGTAAGA ATAAATAATA GAGGAATGCC TAAAATTATA
 190351 GCAGCGGTTT CAGAACTTAC AGCAATTCCT ATAATTAGAA TAAAAGCGAC
 190401 AAGGTCCAGA CCTTTGAGCT TATGAGATGC GAGTTGCGTG GGCATGCTTT
 190451 CGCTGCTGAT CAGCATGGTT AGATTACGT TAGTCAAATT GGGTCTGTG
 190501 GTGGACATAA AACTTTTTTA ATAAAAACA AAAAGTTTAA AAAAGATTCT
 190551 TTTTATGTG AAGTTATTTA TATTTTAAAT AGAAGTTGTT TATTTAAAT
 190601 AATAAATAGA CAACATTTTC ACTTTAAAAA GTATTGGATG CGGCTTAGAG
 190651 CCAAGAATCT AGGGGGTGCT TGTGATCAGG ATAAATTGGC TGTGTTTTTA
 190701 GAAAGCTGTT CGATTGGAAG CACGACGCT TTCCGACATA GACTAGGGAG
 190751 TCGTTAGATC GAAACGGGGT GGAATGATGG GAGGCTGGTC CTTGTCTGTA
 190801 AGGATGATAA TTTTTTGCTC TTCCTGGTTG TCTTGTCCC ATCCAAAATC
 190851 TTGAAGAGAT GTGATGAGTG CCAATTTTAT AGCCTCGATC CATTCTGCTC
 190901 TTGTTTTCCC GAGGTTTAGA AGTCTTGATG GAGCAAATAG ATTACATCCA
 190951 ATGAGGGGGA GTTGAATCAC ATCAACGCCT ATGATTTCAG CTTCTTGGA
 191001 CAGGTTATGA AACGCTTCT TGCTTACTTC AAATGCTTGC TTAGGATCGT
 191051 TGTTACTT AGCAGCTTCG GGACCAAGCA GTTGTGCTAA GAAGTGTGCT
 191101 TTGCCTGGGA CATGGTCGTT TGAGGTGTGA TCACTATTTT CCCATTTTGC
 191151 TGAGCGGCAT TCTCCTGGCT GTAGTTGGGA TCCAGAACGA GAGTGCGCTC
 191201 TAGGGAGCCT AGATGCGTTC CAACACTGTA GACTTGTAGC CAGGGATAGA
 191251 GCTTTATTCG TTCCCCCTCC TTCTCGGGAG ATGTTCTCGT TTGCTGCGTT
 191301 AACAATCATC ACCCTGCCTT GAGTTTTGAT TCTTGGAAT GCAATATCTC
 191351 CTGAGGTAGA TATAAAGATA AGCTTCGAGT CTTTAAGCTT CCAAATAAAA
 191401 CAGGGCTGCT GAGGTGTTTC TGTAGTAAAC GATGGGTCTA CAGCGGCTAA
 191451 GCTCGGAAGA AGATCCCAGT TTTGTTCAAG AAGAGCCTCA TAGCCAGGAG
 191501 TGACCGTGAT AGCGGTAAGT TCTTCTGGGG GTGTGGTTAG ATCAGCAAGA
 191551 GAGACTTCGG GGAGTGATTC GGGATCGGGA AGTGGATCGA TCTGCAAAGG
 191601 GAGGTCCTGT TGCTCAGGAG GCTGGGGAGA GGGAGACAGA GAACTTTGCT

191651 CAGATTCGGG TTCGATTGA GGAAGAATCT CGTATAATTT AGGTTTCTTT
191701 AAAAGGGAGT AAAGTGAGAC CGCAGCAATT GCTAATGTTA TGACAACTAC
191751 GGCAGAAAAA ATAGGCATAG CCAAAATGCC AGCAGCAATG CTAAAGCCCA
191801 AAGCGCACGC TAGGGCTAGT GTGAGGGTGA CAATTTTCAG TATTTTCTCG
191851 ATTCTATCTG TACGGTTGAG AGGGAGTCTA ATCGGATAGG AATGTTTCGC
191901 AGGAGTTCTG TAGAGACTGG CGTCTGTATA AGAGGGTAGG GGATTAGAAT
191951 CTGTCATAAT ATTTTATTAA TAGATCAATT TTCAACTATT ATATTACTAA
192001 TTTGTATTTT TATTAGATTT TTTTATAAAA CAAAATTAGT TTATTAATGC
192051 ATCCTATTGA AATAGATCTT TTTAGTTAAA AGGGACCATA AGTAGCTGTT
192101 TGTGGTCTGT AAGGATTATA GTCATGGGAG TTGAGGGGTG CTGCGCAGCA
192151 AAGGAGCGAA GAGCTTCCAT AAGACCTTTC TTTATATCAT TTACCCACTA
192201 CGTACGGATG TGGTAAAGCT TATATGCACT GCTATTAGGC TTTGTTTGGT
192251 TTACGGGTTC TAGTTCCAGC TTTCTCCAG GTGAGTATAT AGAGGAAGAG
192301 ATCAGAGGCA CTTGGACCAC AGTGGCTTGG TTATTGAGAG CTTTCATCAA
192351 ACAGTTCAAA TAGGCTTTCT TAATAACATT GCTTAATTTTC TCAGGATGTG
192401 CTTTCAATTC TCCTTCATAT TTAGGACCAA GAAGTTGTGC TAAGAAATGT
192451 GCTTCTCCTG GGTTCGTATC ATTTATTCGT CCAGTCTCTA TTGATCCAGG
192501 GTGCTGAGCG GCATTCACCC ACAGATAATC CTTTGCCAGT GTTTATTTTT
192551 CCCCCAGATG TTCTCGTATT GTTCCAACAA GTAGGGTGTG TGGCTGCTGA
192601 GAGAGCAGCA TTGGTTCCGG CTCCACCAGA TTGCATGTTC GAATTCGCTG
192651 CGTTAACAAT CATGACTCTT CCTGAGGTTT TCAGGCGTGG TTTAGCGATG
192701 TCTCCTGTAG TGGCTACTAA GGTAAATTGGG GCTCCTTGAT GTTCCCAGAC
192751 ATAGTATCTT TGATTAGGAT CTTGGAGAGT CCAGGATATA TTGATTTCTG
192801 AGAGAGTATT GACTAGGGTC CATTCATTTTC TCAGCAGCTT TTGGTAATCC
192851 AGAGGGACTG AGGGTTGCAC TTGAGCTGCA ATATCCTTGG GGACATGTTT
192901 TTGGGGCCCA AGGAGCTCTT CTGTCGGTTT GGATGGCTCT CTTCTTCTTA
192951 AAAGAAGACA AGAGAGTACG ACTGCTGCAA GGAGAGCAGC ACCAGTGGCT

193001 ATAGCCATGA GAGGCATACC AAGAACTCCA GCAACAACTG CAGTCCCTAC
 193051 AGCTATGGCT AAAGCAAGAA TAAGAGCTGT GAGTTTTATT ATTTTAGCAA
 193101 TCACTTCTTT CTTAGTAAGC ACGGGTGATT CTGAGATCAA TCTCTCATCT
 193151 ACAGACAAAT GTGATTGATT CTCGTGTACA TCTGGGATTG AAGTGGTTTT
 193201 AGTCATAAAA TATCTATTTA AAGGAAAAGA TTATATCTTA ATTTGTTGTT
 193251 TTTTACAAAA AATCTTTTTTC TTTGATAATA AAGATTTTTA TAAATTCTTT
 193301 TTAATTTTCAT TAATTTTAAT GAGTTAAGGC GTAGCAGGCG CACCTTATTG
 193351 TTTAGTTCAT AAGAGAACTC TAAGTGGGTG CCACACAGAG TTTTTTCTCC
 193401 ACTTGGTTCA AAAGGGTGGC GTTTGTGGTT TTATGGGGAA CGATTCTACC
 193451 CCAGCCATGG GATTCAAAGA CTTCTTAGAT CGACAATACT ACTTGAGATT
 193501 CTAATGGGGG GTGAAGACAT GTGTATTTCA TGGGCATGGG TAACGGAGGA
 193551 CATCAGGTTT ACTTTCGAAG GCCCATATAT CATCTTCAGA GGTCAATAAA
 193601 AGGGGGGGGT AGGGAGAACG ATGTTTTACA ACAGTGTAGA AATTAACAT
 193651 TTGTGAAATA TTTTCATTTA ATAACAATTT AATGAGTTTT AATTTGAAGA
 193701 CACAGGCAGT GGCTTATCTT TTGTCTATAA GTCGGTACAG TCAGACTTAT
 193751 AAGAATTGCC GCTTTTTGGG TCTTCATAAT CTCTATCTTA AACAAAAAGG
 193801 AGGTCAATCA CAACTGGGT GAGTTGTAGA GGTCTCAATA AGAAAGTTTG
 193851 TAAGGTTAGG AGAGCTATAA TTGCGCCTGG GATTCTGCAT AGACGATACC
 193901 TGAAAAGATA GTGCCTAGGG TCGGTGTTTG CTGATGAGAA TGCCAAGAAG
 193951 CATAGTCATC AACCGTCTTA ACATATTTAT AAGTTTTCTT AGCAGCGAAT
 194001 GGCACAGCTG CTAGAAAGCA AAACAAGATG ATAACAGCCA CCAATAGAAT
 194051 TGTTCCCTGAG ACAGCAAGGA CTGCAATGGC AGCTGGATGT CCAGTTTCTA
 194101 GAGTGAATTG CAGAGCAGGG CGGATTGCCA ATAACTAAG AACTATGAGG
 194151 CAGCATCCGA TCACTAGGAG AACTTCCAAT ATGACAGAAG TGATTGTCCG
 194201 TGGAGAAGAC TCAGCTTGCT GCTTCCCTT GAGTTCTATG AGTTTCTGAT
 194251 TTTGTGTGTT TCGCTATTG ATTAAGGGAG CGGCTGGACT TTCAAAAAGA
 194301 ACGTCCGTAG CAGATACTGG AAGATATCCC ATTTATTTTA ATATCCTTTA

194351 ATTAAAATTT GCGCAGATAT TATACCTTAA TTTAATTTTT ATTAGAGAAT
194401 AAAATTCTAT ATTTTCTTTT GTGTAATATT TTTTATAAA AACATGAGTT
194451 TTTAATTTAA CTTAGAAGAG AAGTGTTGCA ATATCTTTTA TTTAAACAAG
194501 AGTACCTAAG AATTCGTGCG TAAAAATATTA CGTATTTGAT TTGTATAAGT
194551 CGTGAGAGTG TTTTTTCTAT TATTGAGGGG CTATTTTGAT TTTTCTTGAA
194601 TTGCGTCATG GAAAGATCTT GGAGGCTTTA GGAGCCTTTT TGTCCCTCAT
194651 CATCTCGAGA TAACCAAACCT GAAAAGATGC GTACCAAGAG TTTGTTTTTT
194701 CGTTCGCGAA TCTACTGTAC CCTTGATTTT AGGATAGCCG TTTTGAGTGA
194751 GATAGAGATC CTAAGAGAGA TCCTCAATAC TTTCTTGAGG TAGTTGGGCT
194801 AAGGATCTTT CAACACAAGC AATAGTTAGC GGATATTCTG GGTGCAGCTC
194851 GGCCATGTAT TTTACTGCTG TAATCATAGC CAAAGATTGT TGCATTCTTG
194901 ATTGCGGGGG CAATTCTCTA GAATAGAGAT GCTCCCTCAA GCTGTATAAA
194951 GGGATTGGA TAATTTGGGT GTTCGTATTG ATCGCAGTCT GGATATAAGT
195001 GACATAGAGT GCATAGTAGT ATTGGTAGAT GGCTTCAGGA GTTGTGTTGA
195051 TTTGTGGGAG GAAGACGCGG CCTCGGAAAT CGGGTTGCCA GTATAAACGT
195101 TCTCCTGAAT CTCCCGCAGG GAGTTCTGTA TTTATTTCTA CCCTGGTTTC
195151 CTGAATTGCC TCATTGAGCA AGGATTTTGG GGAGGGAGTG GTATCTACGT
195201 AGGAGGATAT AGGCGCATTT TCGGAGGTTT CTTCAATTAA CAGAGTCTTT
195251 CCTAAATTTT CGGAAGCTGA AATTAAGAGA TGTCTGGAT TGCCTCCCGG
195301 TAGGAGGGAT ACTAGGGAAA ACAGTGAATG TTTAGCTTTC CATATCTCGT
195351 AAGGGTTATC TGTGAGTAAT GCAGAGGGTG CGGGGAGATT CTGTAGGTCC
195401 TTGGGTTGGT GTTCCCAGTT TTCTTTAGAA CGATTGATTG CGATAGAGAT
195451 AGAAGAATAC TGTAGGTTGG TGGAGACCCT TTTTGGTTTT ATTTCTTTGT
195501 GTTTGCCACA AGCACGCGGT TTTAAGAGGG TGGTTGCTCG TTTATAAAGG
195551 ATAAAGTTGC TAAGAACAAC AGCAATAAAT GAAATTCCTG TGAGGAAATA
195601 GATAACAGGA ATATTTAAAA GCAAACCAAT AAGTAGGGTC CCTATAGTAA
195651 CTAAGACAGC AAAGATTGCA GTAATCAATA GAATTCGAGA CTGAATCTTA

195701 TCAAAGTTGC TGCAGGATTT TTTTCTAGTG TCTACTAGGG GATGATCTAC
 195751 AGTTAAAGAT ATGGGTGAGA TTGTAGCCAT AATTAAGGAA TTCACCTGAG
 195801 TTCCTTTAAA TTATATCTCT TTTATTGATA TTTAAATAAT TAAAAACATT
 195851 TCTGCATTTA AAAATATTTT TTAATTTGTC ATCTTTCTAA GTAAATATAA
 195901 AAAATAATAA GTTTATTTAG TTAAATTTT TTTGAAAACA AGAAAGATAG
 195951 GGGAAAGGCT CGCGAGTAAA TCAAGCCTTT CCTGAATGCA AAGAGTCGCA
 196001 TAACTACTTA GAAAAGAGTT CGTCAGAATA GAAGTTGTAT TAATATAAAA
 196051 TTTTCTATTT CTAAGATTAG AAAGTAACTT TGACACAGCC ACCTTTGATG
 196101 CGGCACTGAC AAGCAAGACG TTCGTTAGAG TCTTCGGGTT CTCCTAGAAA
 196151 ATCGTATTCT GGTTCGTAA ACTCAGAAAG ATTCTCACGT CCTTCTAAGA
 196201 CCTCTATCAC ACAAGTTCCA CAGACACCTT CTGTACAAGC AAAGGGAATG
 196251 CCCATGGATT CACAAGGCTC TCGCATCTCA CTATTGTCTT CTAAGTCGAA
 196301 CTCTTGTGTG TCATCATCAG AGGTAATGAC TAGCTTGGCC ATGGAGTTTT
 196351 CCTTCTACGT ATTGTCGGAT AGATTAAAAAT GAAGTTCGGA GTAGAGGGAT
 196401 TCGAACCCCC GACCTATTGC TCCCAAAGCA ACCGCGCTAA CCAAGCTGCG
 196451 CTATACTCCG TAAAATAAAA TTTTCGAGTA AAAGAGATCA ACGTTAGCAT
 196501 AGAAGGAAAT AGTCGACAAG ATCAAAGATG CTTCATAACG TCTCTTTAGA
 196551 GTTTTACTTG CAGATATATG GAAGAGGGAA GTATGAGAAA AGGCAGGGAT
 196601 TCTCTAGGAG GCTGTTTTTG TGTCTGGGAA GAAAGATGGT GTAAGGGGAA
 196651 TGATCTTTGT CCCTCTTAGC ATCCTAGTAC TAATCTTTTT ACCTCTTCCT
 196701 CAGATCCTTC TTGATTTTGG ATTGTGTATT AGTTTTCAT TGTCTTTACT
 196751 AACGGTCTGT TGGGTCTTTA CCTTAAATTC AAGCAATTCA GCGAAGCTTT
 196801 TTCCTCCATT TTTCTTATAT CTTTGCCTAT TCGGTTGGG ATTGAATCTT
 196851 GCATCAACAC GATGGATTGT CTCTTCAGGA ACCGCCTCTT CTCTGATTGT
 196901 TTCTTTAGGC AGTTTCTTCT CTTTAGGAAG TCTATGGGCA GCAACGTTTG
 196951 CGTGCCTCCT TCTTTTCTTT GTGAACTTTT TGATGGTTTC AAAGGGTTTCG
 197001 GAAAGAATCG CAGAGGTCCG TTCGCGGTTT TTCTTAGAGG CTCTTCCAGC

197051 AAAACAGATG GCTTTAGATT CTGATCTTGT TTCTGGAAGA GCTTCTTATA
 197101 AGGCTGTCAA AAAACAAAAA AATGCCCTTA TAGAAGAAGG GGATTTCTTC
 197151 TCTGCCATGG AGGGGGTCTT TCGTTTTGTT AAAGGGGATG CAATTATTAG
 197201 TTGTATCCTT TTA CTCTGTA ACGTAGTTTC TGTAACCTGT CTTTATTATA
 197251 CTTGGGGTTA TGCTCTTGAG CAGATGTGGT TTACAGTTTT AGGAGATGCT
 197301 TTAGTGAGTC AAGTACCTGC TTTACTTACT TCGTGTGCTG CAGCCACTCT
 197351 TATTAGTAAA ATCGATAAGG AAGAGAGCCT TTAAATTAC CTGTTTGAAT
 197401 ACTACAAACA GTTGCGTCAG CATTTCAGGG TGGTGTGCTT ATTGATCTTT
 197451 TCTTTGTGCT GCATTCCCAG TTCTCCAAAA TTCCCTATCG TTTTGCTCGC
 197501 GAGTCTTTTA TGGTTGGCGT ATCGAAAAGA AGAGCCTGCA TCAGAAGATT
 197551 CTTGTATAGA ACGTGCCTTC TCTTATGTTG AGGGGGCCTG CCCTAAGGAA
 197601 CAAGAATCAC AGTTCTATCA AGTATATCGT GCAGCATCCG AAGAAGTATT
 197651 TGAAGATTTA GGAGTTAGAT TGCCTGTGCT TACTTCTCTA CGTATTGAAG
 197701 AGCGTCCTTG GCTCCGAGTA TTTGGCCAGA ATGTATACTT AGATGAAATG
 197751 ACTCCAGAGG CTGTGCTTCC TTTCCCTAGA AACATCGCTC ATGAGGCTCT
 197801 CAATGCCGAG GTAGTTCAAA AGTACCTTGA GGAATCAGAG AGAGTGTTTG
 197851 GCATCGCTGT TGAAGACATC GTTCCTAAGA AAATCTCTTT AAGCTCTCTT
 197901 GTAGTTCTTT CTCGCCTCCT TGTTAGAGAA AGGGTATCGC TTAAGCTTTT
 197951 CCCAAAGATT CTAGAGGCCG TTGCGGTATA CCAAATTCT GGAGACAGCT
 198001 TGGAGATCCT TGC GGAAAAA GTGCGAAAGT CTCTCGGATA TTGGATTGGG
 198051 AGAAGTCTCT GGGATCAGAA ACAAACCTT GAGGTAATTA CCATAGATTT
 198101 TCATGTTGAA GAATTGATAA ACAGCTCATA CTCAAAGTCT AATCCTGTAA
 198151 TGCAAGAGAA TGTGATCCGT CGAGTAGACA GTCTTTTAGA ACGGTCGGTA
 198201 TTAAAGATT TTCGAGCCAT AGTTACGAGC TGTGAAACAC GATTTGAGAT
 198251 GAAAAAATG CTCGACCCAC ATTTCCCTGA TCTTTTGGTT TTATCTCATG
 198301 ATGAGCTTCC TAAAGAAATC CCTATTTCTT TCTTAGGGAT CGTTTCAGAT
 198351 GAGGTTTTAG TTCCTTAATT TAATTTAATC TTCTGTAAGC ACTATTACTT

198401 TCGTATTTTT TTTGTTGGTG TAAATATTGT TTAAAAATTT TTTTGAATT
 198451 AATCTAATAA ATAAGTAGAT AAAAAAAAAAT TTGTGAAAAC ACAGCAAAC
 198501 CAAAACATCA TAGAGGTTTG GAACTTCTAC TGGGAGACTC AGGAAATAGA
 198551 GTATCGCGAT AGCTTAATTG AGTTCTATTT GCCTTTAGTA AAAAGTGTGG
 198601 TTCATCGTTT GATTTTCAGG ATGCCTTCCC ATGTAAAGAC CGAGGATTTG
 198651 TATGCTTCGG GTGTTGAAGG TCTCGTCCGT GCGGTGGAAC GTTATAATCC
 198701 TGAGAGAAGT CGTCGTTTTG AAGGTTATGC GGTATTTCTG ATTAAGGCTG
 198751 CCATTATTGA TGATCTGCGT AAGCAAGACT GGGTTCCTCG TAGTGTCCAT
 198801 CAAAAGCGA ATAAATTGTC AGGAGCTATG GATTCTCTTC GCCAGTCTTT
 198851 AGGCAAGGAA CCCACGGATC TTGAACTGTG TGAGTATCTC AATATTTTCGC
 198901 AACAAGAGCT TTCGGGATGG TTTGTATCTG CCCGTCCTGC ATTAATCGTG
 198951 TCTCTGAATG AAGAGTGGCC TTCACAAAGT GATGAAGGAG CCGGAATGGC
 199001 TCTTGAAGAG AGAATCCCCG ATGAACGTGC CGAGACAGGG TACGATGTTG
 199051 TAGATAAACA AGAATTTTCT TTATGTTTAG CCAATGCGAT TCAGGAACTT
 199101 GAGGAAAAGG AACGCAAGGT CATGGCCCTG TACTACTATG AAGAACTTGT
 199151 CCTTAAGGAA ATCGGTAAGG TCCTTGGGGT AAGTGAGTCT CGCGTCTCTC
 199201 AAATTCACTC TAAAGCATTG CTTAAGCTTC GCGCAGCACT CTCTGCATTT
 199251 CGATAAATAC AGTTCTCAGG TTTAAGAGC AGTCCTAGAG CTAGGAGAAG
 199301 CCCTCCTACG ACACAGAGTA ATCCGTAAAG AATTTGTTTA GTCAACGCCA
 199351 TAACAATGCA GCCAGCAATT GTGAAGATTG CTCCTACAAT CAAGAGAAGA
 199401 ATGCGGGGCA GCCATTCTTT AATACGTTTC GCTAGGGTAG AGGTAGGAAG
 199451 GGGTGCCTTG CTAGTTGTGG CATTTAGACT AGTTGCGGCG TTTAGCATCG
 199501 ACGGAACCCC AATTGGAGAT GTAGACATAG CGACCTAATC TTTAGAGAAG
 199551 ATAGAAGTTA TGGTATCTCA GCAAAGGAAT TATTCTCAGC GTACAATCTT
 199601 TTCTTTATTG TTAGATAGGT GGTGCCTTGC ATATAAGCTT TAGGATCGAT
 199651 AAGATAGAGC TTGGCAATTC TTAATTATGT ACGATCACTC ATGCAATCCT
 199701 GGTTACAATC TTTACAAGAG CGAAATATTT TAGAGAATTT TACCGCAGGT

199751 TTGGAATCCG TAGAGGGACC TATCGCCGCT TATTTAGGAT TTGATCCTAC
 199801 CGCACCTGCT CTACATATTG GTCATTGGAT TGGGATTTGT TTCTTGAAGA
 199851 GACTCGCTGC TCTGGGGATT ACCCCCATAG CTTTAGTCGG GGGAGCCACA
 199901 GGTATGGTTG GAGATCCCTC AGGGAAACAG AGCGAGAGAT CGTTACTTCA
 199951 GACAAGTGAA GTTTTTGATA ACAGTCAAAA GATCACGGCG TGTCTCCAGC
 200001 GCTATCTTCC CGGGGTGACT CTTGTAAATA ATGCAGACTG GTTGCAGGAG
 200051 ATCTCCCTGA TTGATTTCTT AGGGGATATA GGAAAACACT TTCGTTTAGG
 200101 CCAAATGCTA GTGAAAGATA CAATAAAGCA GCGGGTGCAT TCTGATGAAG
 200151 GAATTAGCTA TACCGAGTTT AGCTATTTAA TCCTGCAATC CTATGATTTT
 200201 TATCACTTAT TAAAAATTA TGGCACGATC TTGCAGTGCG GTGGTAGCGA
 200251 TCAGTGGGGG AATATTACTT CAGGAATCGA TTTTATTCGC CGTAAAGGGT
 200301 TGGGTCAGGC CTACGGCCTT ACCTATCCTT TATTAACGAA TGCTCAGGGG
 200351 AAAAAAATAG GGAAAACAGA GTCGGGAACT GTATGGCTCG ATTCAGATTT
 200401 AACCTCTCCT TTTGAGCTGT ACCAATACTT ACTCCGTTTG CCCGATGATA
 200451 CCATCCCTAA AATTGCTCGT ACGTTAACTT TATTGAGCAA TGAAGAAATT
 200501 CAAGATATTG ATAGGCGTGT ACAGACGGAT CCAGTTGCAG TGAAGGAATT
 200551 TGTAGCCCAA GATATCTTAA GTGCTATTCA TGGAGATCTA GGGCTTGAAG
 200601 AGGCTCTTTC TGTAACTCGT AGCATGCATC CAGGGAATCT TTCATCCTTA
 200651 TCGGAAAAAG ATTTTCATGA ATTGTTTGCA GGAGGGATGG GGGCCTCATT
 200701 GGATAAATCC GAGGTGTTAG GGAAACGTTG GTTAGACCTA TTTCTTGTTT
 200751 TGGGACTATG TAAATCTAAA GGGGAAATTC GAAGGCTAAT TGAACAAAAA
 200801 GGGGTATATA TTAATAATGT GCCCATCGCT AATGAGCATA GTGTTTGTGA
 200851 AGAACAAGAC ATCTGTTATG GTCACATATG GTTGTTGGCT CAAGGTAAAA
 200901 AACGAAAGCT TGTTCTATAT TTAAATTAGT TTAAGGAGGC TAGGTAGCTT
 200951 TGCAAACGAA TATTGGTCTT ATTGGCTTAG CTGTCATGGG GAAAAATCTT
 201001 GTCTTAAACA TGATAGATCA TGGTTTTTCT GTCTCTGTCT ATAATCGGAC
 201051 CCCAGAGAAA ACGCGGGACT TCTTGAAAGA ATACCCTAAC CACCGAGAGC

201101 TTGTAGGGTT TGAATCTTTA GAAGATTTTG TGAATTCATT GGAGAGACCA
 201151 CGAAAGATCA TGTTGATGAT TCAAGCAGGG AAACCTGTGG ATCAGAGCAT
 201201 TCATGCGTTA CTGCCTTTTC TAGAACCCGG CGATGTGATT ATCGATGGGG
 201251 GGAATAGCTA TTTTAAAGAT TCCGAACGAC GATGTAAAGA GTTGCAAGAA
 201301 AAGGGGATTC TCTTCTTAGG CGTGGGGATT TCTGGAGGAG AAGAAGGTGC
 201351 ACGTCACGGC CCATCAATTA TGCCTGGAGG AAATCCTGAG GCGTGGCCAT
 201401 TAGTGGCTCC TATTTTTTCAA TCAATAGCAG CAAAAGTACA GGGCCGTCCC
 201451 TGCTGTTCTT GGGTAGGAAC TGGCGGTGCA GGCCACTATG TAAAGGCTGT
 201501 TCACAATGGT ATAGAATACG GCGATATCCA GTTGATATGC GAAGCTTACG
 201551 GTATCTTAAG AGATTTTCCTA AAGCTCTCCG CAACTGCCGT TGCTACAATT
 201601 TTGAAAGAGT GGAATACTCT AGAGTTGGAA AGCTATCTAA TTCGTATTGC
 201651 TTCTGAAGTC CTAGCATTGA AAGATCCGGA AGGAATCCCT GTTATTGATA
 201701 CGATTTTAGA TGTCGTGGGC CAAAAGGTA CAGGAAAGTG GACCGCAATC
 201751 GATGCTTTAA ATTCTGGAGT TCCCCTTTCC TTAATCATAG GAGCTGTTCT
 201801 TGCTCGTTTC CTTTCTTCTT GGAAAGAGAT ACGCGAGCAA GCTGCCCCGT
 201851 ATTATCCAGG AACCCCCTTA ATATTTGAAA TGCCCCATGA TCCCTCGGTA
 201901 TTCATACAAG ATGTCTTTCA TGCTTTATAC GCTTCCAAGA TCATCAGCTA
 201951 TGCTCAGGGA TTCATGCTTT TAGGAGAAGC TTCAAAAGAA TATAATTGGG
 202001 GATTAGACCT AGGAGAAATT GCTTTGATGT GGCGCGGGGG ATGCATTATT
 202051 CAAAGTGCAT TTTTAGATGT TATACATAAA GGATTGCTG CCAACCCAGA
 202101 GAATACCTCG CTCATCTTCC AAGAATATTT CCGTGGAGCA TTACGCCATG
 202151 CGGAGATGGG ATGGCGTAGA ACAGTAGTGA CTGCAATTGG TGCAGGGCTA
 202201 CCTATTCCTT GTTTAGCAGC AGCAATCAGG TTTTATGATG GCTATCGTAC
 202251 AGCAAGCTCT TCAATGTCGT TAGCTCAAGG ACTGCGAGAT TATTTTGGAG
 202301 CTCATACCTA CGAGCGTAAC GATCGCCCTC GAGGAGAGTT CTATCATACC
 202351 GATTGGGTGC ACACGAAAAC TACAGAAAGA GTGAAGTAAA AATAAAAAAT
 202401 CTCGAGAAGA CCCTAGGTAG CTCGAGATTT AGTTCACCAC TACCGAATTT

202451 TCAATTGTAA GCAGTGTGCT GATTTTAAGC GTCAATGTTA ATCTAATTTT
 202501 AGAACTTCAA TGAAAGCTGT ATTGGGAATG GAAACTTTTC CAAATTCCTT
 202551 CATACGTTTT TTTCCTTTCT TTTGCTTTTC CCACAGCTTG CGTTTCCTAG
 202601 TAATATCTCC GCCATAACAC TTTGCGGTCA CGTTCTTAGA AAGCGCACGA
 202651 ATCGTTTCTC TGGCAATGAC TTTTTTGTTA ATGGCAGCTT GGATGGGAAT
 202701 CTTGAAGAGT TGTGTGGAA TCACGTCCAC AAGCTTTTCG CAGATACTTC
 202751 TTCCACGAGA TTCTGCTTTA TCTCTATGGA CTAAACAAGA AAAAGCATCT
 202801 ATGGGCTCCT CGTTAATAAG AACCTCTAAT TTGATGATCG ATCCCTTACG
 202851 GTAATCCCCA AGACGGTAGT CAAAGGATCC ATAACCTTTA GTTACTGACT
 202901 TCAGCTTGTC ATTGAAATCC GAGACAATCT CATTTAAAGG GAGTTCGTAA
 202951 GCAAGAACTA GACGGTGCTG ATCTAGCATT TCTGTTTTTA CGCAGATCCC
 203001 ACGTTTATCT AAACAGAGGT TCATAATGTT GCTCAGATAT TCTTGAGGGG
 203051 TGATAATATT CACATGAACC CAAGGCTCTT CCACATGCTC GATGATCGCA
 203101 GGATCCGGAT ATCCTGAGGG GTTATCAATA TCTAGAACTT TCCCGTTTTT
 203151 TAAGACGACT TTATAGATGA CACTTGAGAGC CGTTGCAATA ATATCTAAGT
 203201 CAAATTCCTG AATGATTCTT TCAAAGATAA TCTCAAGATG AAGAAGTCCT
 203251 AAGAAGCCAC AACGAAAACC AAAGCCTAAA GAGTGACTGC TTTCTTGTTT
 203301 TATAGTTAAA GCAGAATCAT TGAGCTGTAG TCTTCCTAAA GCATCTTTCA
 203351 AAGTATCAAA ATCAGAAGAA TCTATAGGAT AAATTCCAGC AAAAATACC
 203401 GGATTGATCT CTTTGAAGCC TTCCAAAGGA GTTTTTGCAG GATGTTTTGT
 203451 TTTCGTGACT GTATCGCCGA TCTTCACATC CTTCACTTTT TTGAGATTGG
 203501 CAATAAAAAA ACCCACCTGA CCAGGGCGTA AGGAACCTTC TATAAATGTT
 203551 GCTTTAGGGA GAAAGGCCCC TATACCTAAG ACTTCAAACG AGGAGCCTTT
 203601 AGCCGCCATA AAAGTAATGC GGTCTCCTTT TTTTAATTCC CCGCTAATAA
 203651 TGCGTACGTA GACCATAATG CCAACGTAAG GGTGATAATG AGAATCAAAG
 203701 ACTAAAGCTT TAAGCTCTGT TTCTGCAGGT GCTTTTGGAG GAGGAACAAG
 203751 ATCGATAATT GCTTTCAGGA TTGCAGGGAT CCCCTGACCT GTTTTTGCAG

203801 AACAGGCAAT AATGTTTCGTA GTGTCTAGGC CTATATAATC TTCAATCTGT
 203851 TGAGCAATTC TCACGGGATC AGCGGCAGGT AGATCAATCT TGTTTAATAC
 203901 AGGAATGATC TCTAAATCTC TTTCAAGGGC CAGGTAGACA TTAGCAAGAC
 203951 TTTGTGCCTG CACCCCCTGG GCGGCATCTA CAATAAGTAA GGCGCCCTCA
 204001 CATGCAGATA GAGATCGAGA GACTTCATAC GAAAAGTCCA CGTGACCAGG
 204051 GGTATCAATC AGGTTTCAGTT GATACACCTC TCCTTCATAT AGATACGTCA
 204101 TGGTGACAGG ATGAGCTTTA ATTGTAATGC CACGCTCTCT TTCAAGATCC
 204151 ATGGAATCTA AGAGCTGCTC ACGCATCTCC CGTTCTTCTA CTGTGCTCGT
 204201 ACTTTCTAAA AGGCGATCAG CAATTGTAGA CTTCCCGTGA TCAATATGCG
 204251 CTATGATTGA AAAATTGCGA ATGTTCTCTA TCTTATATTC TTTCAAAATG
 204301 TACTGTAGTG TTATCTAGGT TTATTCCTGG TTTCATAAAG CTGCATTGAG
 204351 AAGGCTCCTT CAGACAAAGA CCAGCTTCTC AAATACTTTT GCGGACTGCA
 204401 TCTCTTTATT AGCAAAATGA TCATGCGGTC ATCTTTAGCT TTGGATAGAT
 204451 GAAAGTATAA TAGATGAGAT GTAGAAACCA CAAGGGTTTA AAGTCGAATC
 204501 TGATTTGAGG TAGAGTTCCA GAGTGGCTAG CAGTAGAAAA AACTAAGTGA
 204551 GAAGAAGTGC TCTCCGTTAG TATGAAACTG ATTCCCACTC AGGATTCTAT
 204601 AGAAAGGGAA ACCGATTCTA AAAGAGATAA AAAAATATTT ACCATTTACA
 204651 TATGTTTCATC TAAAGTCCTT GCGGGTCATT TTTTCAGTCA TTTAGACAAG
 204701 CATAATAAAA TTCATGAAAG CATTGGGGTT TGAGATAGTT AAATCGACTC
 204751 GATCCAAGAG TATAAGAGAG GAATGAGTGT TCTTATGTCC GAGATAGTGC
 204801 TCTTCAATGT AGTGAAGAAC TGGAAAGTCA GGAGATCTAT TTGAGTAGGA
 204851 GTTCAGGTCT TCGCAATCGA TTTTTTCAAG CTCTGGAGTC TATGTGAAAG
 204901 AAATTTATAG AAGTAAAATA ATTCCATGAA TTCTAAAATG ATTAGGAAAA
 204951 TAATATAACA CGCTGATACT CAAGTCACAA TGGAGCTTCT ATGGTTAATA
 205001 TACAGCCTGT GTATAGGAAT ACCCAAGTCA ACTATAGTCA GGCTACCCAA
 205051 TTTTCGGTGT GCCAGCCAGC GCTTAGCCTG ATTATCGTTT CTGTTGTTGC
 205101 TGCTGTACTC GCTATTGTAG CTTTGGTATG CAGTCAATCT CTTTTATCCA

205151 TAGAGTTAGG AACTGCTCTT GTTCTAGTTT CTCTTATTCT TTTTGCTTCT
 205201 GCTATGTTTA TGATTTATAA GATGAGACAA GAACCTAAGG AGTTGCTGAT
 205251 CCCTAAGAAA ATCATGGAAC TCATCCAAGA ACATTATCCA AGTATTGTTG
 205301 TTGATTTTAT TAGAGATCAG GAGGTTTCCA TTTATGAGAT ACATCACTTG
 205351 ATCTCTATTC TTAATAAGAC GAATGTTTTC GACAAAGCAC CAGTATATTT
 205401 ACAAGAAAAA CTCTTACAGT TTGGCATTGA GAAGTTCAAA GATGTACATC
 205451 CAAGTAAGCT CCCTAATTTT GAAGAAATTC TTCTACAGCA TTGCCCATTG
 205501 CATTGGTTGG GACGTCTGGT ATATCCCATG GTATCGGATG TCACTCCAGG
 205551 AACCTATGGA TACTATTGGT GTGGTCCTTT AGGACTGTAC GAGAACGCTC
 205601 CCTCTCTTTT TGAACGTCGA TCTCTTCTAT TGTAAAGAA AATTAGCTTT
 205651 GGAGAGTTTG CTCTTTTAGA AGATGGTCTC AAGAAAAACA CGTGGAGTTC
 205701 TTCGGAAGTC GTTCAAATCA GACAAAACCT TTTTACAAGA TATTATGCTG
 205751 ATAAAGAAGA GGTAGATGAA GCAGAGTTAA ACGCTGATTA CGAACAGTTT
 205801 GATTCCCTCC TTCACCTTAT TTTTCTCAC AAGCTCTCTT GAAAGCAAGT
 205851 GCAATTATTT CAATATATGA ATGAGTCCGG ATGGGATTGG CTTTGTGATT
 205901 TTGATTCTCA AGGCGAGGGA TTCCAGTTAT CACGTCTGGT TGGGCTGTTA
 205951 CATTGCTCCT GGGCATTATA CGAAGCAAAA GAGCAATTTT ACCTTCCTGA
 206001 GGTTCCTCTA TTGACCTGGG AAGAACTGAT AGAAATGCAG TTATTAAGCA
 206051 AACCAACAAA ACACGGGGTT GCAAAAGATC TTTGTAATGT ATTTGAAAAA
 206101 CACTTTCAAA GGTTTAGACA GTACCTAGGT TCCTTAGATC TAAATCAAAG
 206151 GTTCGAAAAT ACCTTCTTGA ATTATCCTAA ATACCATTTA GATAGGGAGT
 206201 GAGAAAAAAA TCCTAGGTCA GCTTGGCAGA AATTTTGTAA ATCCTTAAGT
 206251 GTTCGATCTG CATTTTTTTC GGGGATTGTA AAAAGTTTGC CTAATCTGAA
 206301 GGGTAAGAGA GAGCTATTTT TCATGGGATT TTTTATTACA GGAAATCTTT
 206351 GAATTAGAGA TTTTATTTA CACATAATCT ATACTGCCTT CAATAGATCT
 206401 ATATCTAAGG AGTTGGCTAT GAGCATGACG ATCGTTCCAC ATGCTTTATT
 206451 TAAAAATCAT TGCGAGTGTC ATTCTACCTT TCCTTTGAGT TCAAGGACTA

206501 TTGTAAGAAT AGCCATTGCC AGCCTCTTTT GTATAGGTGC ATTAGCAGCT
 206551 TTAGGCTGTT TGGCTCCTCC CGTTTCTTAT ATTGTTGGGA GTGTTTTAGC
 206601 TTTTATTGCC TTTGTCATTC TTTCTTTAGT AATTTTAGCT TTGATTTTTG
 206651 GAGAGAAGAA GCTTCCACCA ACACCAAGAA TCATTCTGA TAGATTTACT
 206701 CACGTGATAG ATGAAGCTTA TGGCCTTTCA ATCTCTGCAT TTGTAAGAGA
 206751 ACAGCAGGTA ACATTAGCCG AGTTTAGACA ATTTTCTACT GCCCTGTTGT
 206801 GTAACATATC TCCTGAAGAG AAAATCAAAC AATTGCCTTC TGAATTGCGA
 206851 AGTAAAGTAG AGAGTTTTGG TATTAGCAGG CTCGCAGGTG ATTTAGAAAA
 206901 GAATAATTGG CCAATATTTG AAGATCTTTT AAGCCAAACC TGCCCGTTAT
 206951 ATTGGCTTCA GAAATTTATA TCAGCAGGAG ATCCACAAGT TTGTAGAGAC
 207001 CTAGGTGTCC CTAGAGAATG TTATGGGTAC TATTGGCTAG GGCCTTTGGG
 207051 ATACAGTACA GCTAAGGCTA CAATTTTTTG TAAAGAGACG CATCATATTC
 207101 TTCAACAATT AACGAAAGAG GACGTTCTTT TATTAAAAAA CAAGGCTCTT
 207151 CAAGAGAAAT GGGATACTGA TGAAGTCAAA GCAATTGTAG AGCGTATCTA
 207201 CACTACCTAT ACGGCACGAG GAACTCTAAA GACCGAAGCA GGGGGACTTA
 207251 CAAAAGAGAC AATCAGTAAG GAATTGCTAT TGTGAGCTT GCATGGCTAT
 207301 TCTTTTGATC AGCTACAGCT GATCACTCAA CTTCTAGAG ATGCTTGGGA
 207351 TTGGCTGTGT TTTGTAGATA ACAGTACCGC ATACAACCTT CAGCTTTGTG
 207401 CTCTTGTAGG AGCTTTGTCA TCCCAAATC TTCTTGACGA ATCTTCTATC
 207451 GATTTTGATG TAAACCTAGG CCTGTATGTG ATTCAGGATC TAAAAGAAGC
 207501 TGTTCAGCA TTTTCTGCTT CTGATGAGCC AAAGAAAGAA CTAGGTAAAT
 207551 TCTTGTTAAG GCATTTGAGT TCAGTTTCTA AGCGATTAGA GAGTGTATTA
 207601 AGACAGGGTC TTCACAGAAT AGCTCTAGAG CATGGAAATG CCAGAGCTAG
 207651 GGTTTATGAC GTCAATTTTG TAACAGGAGC TAGAATTCAT AGGAAGACGA
 207701 GTATCTTCTT TAAAGACTAA ACCAGGTAAC TAGCTTTTTA GTCTCGAAAG
 207751 GGGCTACGAG AGGGAGTAGA ATTTTCTTCA TTCCAAAATA GAAATCTCTA
 207801 ACCTATCAGC AGGGAGGACT CGTCTTCTTC AGGTTGGGAG AAGACGCTTC

207851 TCTTGAGTGG ATGGACTGTT ATAGGTGGTA AGCTTCTAAA GGTGGAATTA
 207901 TGAGTTGCTA AATTCTCAAG GAACATCGCT ATATTGGTTG ACTGACGACG
 207951 GATCGTTTTA AACTGATCTA AACTAAGGAG TGTCAATGAA GAAAGAAATG
 208001 CTTCGCTTTC TTTATGAATT AGGCCTTGTG TGTAGAGAGT ACCAATTAGA
 208051 GAGGCAAATT CCATGTCTTG GGGTCGTCCT GTATGGTCTA AAGTCAAGCA
 208101 ACAGAGGTTT TCCCATAAAT CAGCCGGGAC TGTTTTTATT AAGGACATCT
 208151 GCTTCCACGA GTACCTATGC TTAAAAATAC ACAAGAGCAG GTTATAGAAT
 208201 TCATCCTCTT GAAGAGAGGG TTGAGATAGA TGTTTATAGG ATTGATACGC
 208251 CGCATGATAT TTAGCAAAGA GCTGCTGTTT CATAGTATTT AGATCACTAT
 208301 GATTCAGGTT CCACTCGCCA TTTAATGCTG CATACTTAAG ATGTTGAAAA
 208351 ACATCACGAG TTAAGACTCT AGCTAAAACT AATGTGTGTA GATTTAGAAT
 208401 AGAGGAATTT GTTGGTCTAA AATCTAAAGG AATCAACCAA TAGGTAGGAA
 208451 GCATCTCTGG TTTTATATAA TCACTAAGGT CTTCTTCGAG GCTGCCTCCA
 208501 ATTTTCTTAA CTCGATGGAG CTCTACAACC TTGCTGCCTG CAGAAATAAA
 208551 CTCTCCTAAC CAGAAAAATG GAAATACCTT TTGTAAGATT TTTGGTAAAG
 208601 AGGGGAAGAG GGGACGGGAA GAAGCAGCGC TTTCAAGTCT TTTAAGAGAA
 208651 GCAGTACGGA CTTCATTTTT TAAGCGTGCG ATACCCTCGA ACGTATCTAT
 208701 TTTCTGTTGT AGTTCTTCAG ATACGTTGTA ATTTGTAGAT GATCCAACTT
 208751 CAGAGTGCAA TTGATTAGA AGATCAATAA AACTTATGAG ATCTTTAAGA
 208801 TTTGGTTTAG CTTCTGAAAC AAAATCAGAG ACAAATTTAG GATAGTGC GC
 208851 CCTGATTCTA TTTACGAGTT CTTGGGGGAA TACTTGTCT TTTGATGAAA
 208901 TCATAGGCGT GATTTTTTTT ATTCTAAAA TCACACCAAT CAAGGCTATT
 208951 AATAATCCTA ATCCTAACAA TGCGCCACTT AGAATATAGG AAACAGGAGC
 209001 TGCTACACAC AAGAAAGCTA TCAAAGCTCC GCAGAGTAAG ATGGCACTGA
 209051 TAACAATATG AATAGTGGTT GAATTCTTTA ATTCAAAATA ATAATTACAA
 209101 GAGCGATTAT TTTGAATAAC TGGCGAGGTT ATATTACTCA TACATTTTCT
 209151 AGGCTCATCA ATTTGCGTAT TATATCGTAA ATTTAACAAA AATCCTATTA

209201 ATAGAAATGT TTTTATTTTT AAAATTTTTA TTGAAATTGT TTTGTATTAT
 209251 TGATGAAAAG TTTTCATTGA AGATAGCTTC TAGAAGAAGA ATAACGAGAG
 209301 AGCTTGTTAT GCTTTTTTTTA AAAAAAGGGC GAGCCCCTTT TTGTTCTTTT
 209351 CCCGATTGGG AATAACGTAT TCTTTTATA AAAATTAAAG ATTTTTTTAT
 209401 ATTTCTAAAT CTCAATTATA AAAAGTCTTT CAAACGCAAT TTGCTTGTAG
 209451 GTATTGCAAA TGATCAGACT CCCCTTATAG GGAAGATCGT ATAGGGAGAC
 209501 TGCTTTCTAT TTTTTTCTCT AGTTTGTAGG TTTGGGCGTG GAAAGCATAC
 209551 TGAGGTAAGC TTGATGTCCA TTGTGAACTC GTATTACGAA CAAATTGCCA
 209601 TCGTTTTAAC AGATTTTGAT GGTGCGTTGT ATTTACGCAT ATCTTCTGAA
 209651 CAAGAGCACT CGCTGGGTGC ATAGGACTTT CTTTACTTT CTCTAGTAAA
 209701 ACCTTCAATT CTTTCCACGT CATGAAGTTC ACTGTAGGTT CATAGTTAGA
 209751 GGATACTGGA TCGAACATAT TTGTTTCAGT ATTCAAAAAG CCTCCAAATG
 209801 TTGCCATGGA ACAGTGGCCT CCTGCTTTAT CAAACTGACA CAACATTTTC
 209851 CAATTATCAG GATTTATAAG TTGAATCATC TGAGCCTGTT CCCAAGTGAT
 209901 ACCATGAGAA AAGAAAAGAA ATAAGAATTG TGAGATTCCT TGAACATCCT
 209951 TCCGGA AAAAT CATATTGTGG GGGAGTTCCT TGAATATTTC TTCGCAGGTT
 210001 TTTTTCACAG AAGGAGAATC CCATTGATTC TTAGACGCTT TACTATATAG
 210051 GAACTTATAC TGTGATTCTG AGATTAATGT TAGTAGAGGG CGTGTATAAG
 210101 AGTGGA AAAAT AGTTGTATAT CCTTTATGAA ACGCTAAGGG CCCAAGTAAA
 210151 CCATAAACTT TTTGTGTTTT ATTTAATCCG ATTTCCCCAG CAACAGATTC
 210201 AGTTTTGTCT ATAAAATGGG AGAGCCAGTA TAACGGG CAG TTTTGAAGAA
 210251 GAATCTCTTC GAACTCTGGA AACAGGGTTA AATCTATAGA TTTTAGAATA
 210301 TCGATCCCGA AAGCCTCTGC TTTTTTATGT AAATTCCGAG GCAGGTCTGT
 210351 ACCGCTTTTC CAGCAATTAA TAAATATTTT TAATTCTGTTA ACAGTCAGGG
 210401 AGTGTGTTTT CACGAAATAA AAGACTTCTT TAGGATAGCG ATTGTAAATA
 210451 ATCTTCTGAA GTTCGTTGGG GATGGGCAGA ACCTTTGATT TAGCTAGCAG
 210501 TGCTACGACT AGCGTTATAA TCAAGATTAC GATAGCGGCT AAAGCTAAAG

210551 TTCCTCCAAT AGCATAGCTA ATCGGCGCAG CAAGGAAAAC AAAAGAAAGT
 210601 GCGCTAACAA GAGCTAGAAC AAGCCCAAGA ATGAGTCGGG CAATTGTCCT
 210651 AATTTTTTAAA GAACAAGATC TATGACATTC GCAGTCATTT TTTAAAAATA
 210701 TATGGGGAAC TAGTGTTAAA GGA CTACACT TCATAGTACA ACCAGAATGT
 210751 CTCACGGTTA TTATAACCTT TTGATTGAAA AAAGCGTACG AATTCGTGAA
 210801 ATGGGAGTTA ATAAAAA TCCCTACCCA CTAAAAAGCA GGTAGGGATC
 210851 AACAAGAGTA AGAGAAGCAA CTCTATGAAG AAGCAGGAGC TGAATCTTCT
 210901 TGAGCCACTT CTTGTTCTTT AAGAGCAGAC TCGCTAAGA ATAGTTTGTT
 210951 TAACTTAGTT GCAGAAACCA ACCAAATAGC AATGATGAAA AGAAGAATCA
 211001 CTGCAAGATA AGGGGTCATA GCTCCAATAC TTCCACAGAT AACGAGCAAA
 211051 CCTTGTTGGA TTAAAGCTCC TCCTGATTTT CCGAAGCGGG CAGCAACTAC
 211101 ATCAATAGCA GCCTTACCTT TGACTTTTTG CTCTTGGTCA AGAGGGATAT
 211151 AGGCCATTTC TTTAGTTGAG TCAAAGAGAG CGTATTTTGT GGATTTTCGAA
 211201 AGAATATTCT GTATAGCTCC GACAACCACA GCTAGCATGA GAGGAGTTGT
 211251 ACCGAACATA GCGACCAGCC CAGAAGCTTG GTTTCTAAAG ATAACAAGAG
 211301 CGAAGAAAAC GATACCTGTT AGGAGAACCA TGACAGGAGT GACTAGGGCT
 211351 CCAGTTAACC ATCCAAATTT ACGAATGACG TTACCACCAA CAAATAGCAT
 211401 GATAAGTACG GATACTACGC CAGTCCAGAA GGAGAAGTTC CCCATGAACT
 211451 CACTATAGTC ATTCATATTA GGATATTGCA GTTTCAGCTG ACTTTTCCAA
 211501 GTCACTTCGA TTAAGTTAAT GCAAATACCA TAGGCAATAA CCAAGAGAGC
 211551 TAATAAAAGA ATATAAGGAG ATCTAGCAAG ATAGAGGAAG CTATCTTTCA
 211601 TATTCATTTT AGGTTTAGCA CCTTTTTTCC CCTTTTGCAT TTCTTCTGGA
 211651 TTATAGAAGC GAGGATCGGT CAATACGTTT TTATTGATCC ACCAGTAACT
 211701 GGCCATAAGA ACAAGTCCAG ATACAATAGT CATAGCCATC AAAAGACGTA
 211751 AAGAAATTCC CCAAGGATCT ACACCTTCAG AAACGGAAGC TCTCAACTTT
 211801 GAAGCCCAAA CAATTGCACG ACCAGAAGCT AGTAAAGAAA TATTAGCTCC
 211851 GATACCGAAA AGAGCGTAGA AACGCTTTGC TTCGTGGATT TTTGTAATTT

211901 CATTAGCAAA TCCCCAGAAC ATTAGAGATA GCATGACGCT TCCCCATAGT
 211951 TCAGCAAGTA CATAAAATGC AGCAAATGTC CAGTTTCTTA AGATGGCAAC
 212001 GAGTCCTAGC AATCCTGGAG GTAGGATGGC CTGTAAACGG TCAGCAAATT
 212051 CTGTAGGATG TAAACATCG CGTAGCGGAT AAATTACAGT CGGGAACAGG
 212101 GCAAAGAAAA TTAAAAAGGG CGTTCCTACT GCATAAAATA AGGCCTGCTT
 212151 ACTTAAATA TTAGCTAGCT TTGCATAAAT AAGCATAAAG ATAATAGCAC
 212201 AGGGGACAAC AAGCCAAAAC TTGATGAAAG GTATTGCCTC TGCACCAGAA
 212251 CCAGGAGCTC CCACAATAAG AGTGTCTTTT GTATCGCGTA ACACCGTATA
 212301 GTTAAATGTA ATACAGAAGA ACATTAGGAA CATTGGCAGA ACTTTCTTTA
 212351 GCTCGTGAGT ATGTATCGGC CACAAGAAAG AGCGCAATTT TCCAAAAGGT
 212401 TTTTCTTCGG TTTTGTGCAT ATTTACCCTC TGAAATACTT TTATTTTCTA
 212451 TTCTGATAGT TTTTATTTTC GTTGTGGACT AACTTGAATT GTTTCCATTT
 212501 CGAAAGACCT TATAATATAC CATCTTTTCT CTCTATTGAC AAGAGGAAAG
 212551 GGATACTCCT AAGGGATAAG AGTTCTTAGG CGACTTTATT GCTAAGTCGC
 212601 CTGTAAATT TATTTTTTAA TTAAACGAAA GGAGAGGTAA CTTCAAGGTT
 212651 GGGAGAGCTT ACAGCCTTCA CAGCTTCGAC GTAGGCATTC CATAGCTCTA
 212701 CGCTCACGAT ACCGTTGCGA ATCGAACGAA TTAAGTCCCT TTTTAAAGAA
 212751 GTTAACGGTT TTTTAGGGCT AAAGCGAGCG AGGAGGTCTT TGTCTCCAAC
 212801 TTGTCTTGCC AGTTCTAAAA CGCGTTCAAT ATCTGTTTTA GTAAAGTTAA
 212851 TAAAGTCACG ACGTTTTTTA GAGCCTCTAG GAGCTCTAGG CTTAGGATTA
 212901 ATTGTGCGTG GGGTTTCCGA TTCTAAAATA AATGTTTTTA ACATTTTTAA
 212951 AAGTTGTTTC GGAGCAGCGC TTTTCATCTT TTTTAGAGTG AAATGATGCA
 213001 TATAACCGCC GCTAGGGCCT GGAAGATAGC GACATAAATC GTTTTCTTTG
 213051 TTTCCGCAGA CTCTTTTAAT AGCTTTTCCG ATCAGTTTTT CTATTTCTTC
 213101 TTTTATATGT GATTGTTCTA CGGCCATGAA ACCCACCTTT TATAATTGAT
 213151 GCTTACATAT TTTACAAGAA AGAATGTGGC ATACTTTCAA TTTAATGTTT
 213201 AGAGAAATAA AATTAAAGCG CTGTTTTAAT CAATCAGCGA ACATAAATTA

213251 TATTGCAAGT TCTCCTTTTT GCAATATTTT TCCTCTCAGG ATCTTTTTGT
 213301 TTCTTAACAA AATTCAAAAC ATAAAAATCG TAAAGTAAAG AAGATTTTTT
 213351 CTATAACATT TTGTTTTCTT ATAAAACAAA CGGGTTTTCT AATTTTTTTA
 213401 AAACATATCCT GAACAAAGAT ATTTCTTTCA TTTCTATTAG TTGTATTTCT
 213451 TGCTGGATAA AATATATTTT GAATAGAACT TGTTTTTCTG GTACTTTAAG
 213501 TGGTTTTTAA TGGATCTTAA AAAATGCTTC TAGAGAGATG GATGCGAAAA
 213551 TGGGATATAT ATTTAAAGTG ATGCGTTGGA TTTTCTGTTT CGTGGCATGT
 213601 GGTATAACTT TTGGATGTAC CAATTCTGGG TTTCAGAATG CAAATTCACG
 213651 TCCTTGTATA CTATCCATGA ATCGCATGAT TCATGATTGT GTTGAAAGAG
 213701 TCGTGGGGAA TAGGCTTGCT ACCGCTGTTT TGATCAAAGG ATCCTTAGAC
 213751 CCTCATGCGT ATGAGATGGT TAAAGGGGAT AAGGACAAGA TTGCTGGAAG
 213801 TGCCGTAATT TTTTGTAACG GCCTGGGTCT TGAGCATACA TTAAGTTTGC
 213851 GGAAGCATTT AGAAAATAAT CCCAATAGTG TCAAGTTAGG GGAGCGGTTG
 213901 ATAGCGCGTG GGGCCTTTGT TCCTCTAGAA GAAGACGGTA TTTGCGATCC
 213951 TCATATCTGG ATGGATCTTT CTATTTGGAA GGAAGCTGTC ATAGAAATTA
 214001 CAGAAGTTCT CATTGAAAAG TTCCCTGAAT GGTCTGCTGA ATTTAAAGCA
 214051 AATAGTGAGG AACTTGTTTG TGAAATGTCT ATTTTAGATT CTTGGGCGAA
 214101 ACAATGCTTG AGCACAATTC CTGAAAATTT ACGGTATCTT GTCTCAGGTC
 214151 ATAATGCGTT CAGTTACTTT ACACGTCGCT ATTTAGCTAC TCCTGAAGAA
 214201 GTGGCTTCCG GAGCATGGAG GTCTCGTTGT ATTTCTCCTG AGGGTCTATC
 214251 TCCAGAAGCT CAAATCAGTG TTCGTGATAT TATGGCGGTT GTAGATTATA
 214301 TTAATGAGCA TGATGTCAGT GTGGTTTTCC CTGAGGATAC TCTGAACCAA
 214351 GATGCGTTGA AAAAAATTGT TTCTTCTCTG AAGAAAAGTC ATTTAGTTCTG
 214401 TCTAGCTCAA AAACCATTGT ATAGTGATAA TGTGGACGAC AATTATTTTA
 214451 GCACCTTTAA ACATAATGTC TGCCTTATCA CAGAAGAATT AGGAGGGGTG
 214501 GCTCTTGAAT GTCAAAGATG AGACTTTTTG GTCTGTACAC AACCTTTGTG
 214551 TAAACTATGA GCATGCAGCC GTTCTTTATC ACATATCCTT TTCCTTGGGA

214601 AAGGGGTCAT TAACTGCTAT TTTAGGTCCT AATGGAGCTG GTAAAAGCAC
 214651 TCTCTTAAAG GCTTCCTTAG GCCTGATCAA ACCCTCTTCG GGGACTGTTT
 214701 ATTTTTTTTAA TCAAAAATTT AAGAAGGTGC GTCAGCGCAT AGCCTATATG
 214751 CCTCAGAGAG CTAGCGTGGA TTGGGATTTT CCAATGACTG TCTTAGATTT
 214801 AGCCCTTATG GGGTGTTACA GCTATAAAGG AATGTGGGGG AGAATTTCTT
 214851 CGGATGATCG AAGGGAGGCC TTTCATATTT TAGAAAGAGT TGGTTTGGAA
 214901 TCCGTAGCAG ATAGACAAAT AGGACAGCTC TCAGGAGGAC AGCAACAAAG
 214951 AGCATTTTTTA GCACGTGCTT TGATGCAAAA AGCAGATCTA TATCTTATGG
 215001 ATGAGTTGTT TTCAGCGATT GATATGGCTT CGTTTAAAC ATCTGTAGGG
 215051 GTTTTGCAAG AGCTGCGAGA TCAGGGAAAG ACTATCGTCG TTGTTCATCA
 215101 TGACTTGAGT CATGTGCGTC AACTATTTGA TCATGTGGTT TTATTGAATA
 215151 AGCGTTTGAT TTGTTGTGGC CCTACTGATG AATGTCTGAA TGGAGACACT
 215201 ATTTTCCAAA CGTATGGTTG TGAAATTGAA CTTTTGGAAC AAACCCTGAA
 215251 GCTCTCTCGA GGAAAACAAT TTGGATCGTG CTGATTATGC TCAGTTGTGT
 215301 TTTTCTGAT ACGATTTTCT TATCTAGTTT TTTAGCTGTC ACTTTGATTT
 215351 GTATGACCAC AGCTTTGTGG GGGACAATTC TCTTGATTAG CAAGCAGCCT
 215401 CTTTTAAGCG AAAGTTTATC TCACGCGTCG TATCCAGGAC TTCTAGTTGG
 215451 AGCTTTGATG GCGCAATATG TTTTCTCATT GCAAGCTTCT ATTTTTTGGA
 215501 TTGTGTTGTT TGGGTGTGCT GCTTCGGTAT TTGGTTATGG GATCATTGTT
 215551 TTCTTAGGGA AAGTATGTAA ATTACATAAA GACTCCGCCC TTTGTTTTGT
 215601 TCTTGTTGTA TTCTTTGCTA TCGGAGTGAT TTTAGCCAGT TATGTCAAGG
 215651 AAAGTAGCCC TACGCTATAC AATCGCATTA ACGCCTATCT ATATGGGCAA
 215701 GCAGCCACTT TAGGTTTTCT TGAAGCTACG TTGGCTGCGA TCGTCTTTTG
 215751 TGCTTCGTTA TTTGCTTTAT GGTGGTGGTA TCGACAAATT GTTGTGACTA
 215801 CTTTTGATAA AGATTTTGCT GTTACTTG TGTTAAAGAC TGTTCTTTAT
 215851 GAAGCACTCA GTCTAATTTT TATATCGTTG GTGATCGTAA GTGGAGTTCG
 215901 AAGCGTAGGG ATTGTTTTAA TTTCTGCTAT GTTTGTGGCT CCTTCTTTAG

215951 GTGCTCGTCA GCTTTCGGAT CGTCTAAGTA CAATTCTTAT CCTTTCCTGCA
 216001 TTCTTTGGAG GGATTAGCGG AGCTTTAGGA AGCTATATCT CTGTAGCATT
 216051 CACATGTCGT GCTATTATAG GGCAACAGGC GGTGCCTGTA ACCTTGCCTA
 216101 CGGGACCTTT GGTTGTCATT TGTGCTGGAT TATTGGCCGG TCTATGTTTG
 216151 CTTTTTCTC CAAAATCTGG GTGGGTCATT CGTTTTGTCC GTAGGAAGCA
 216201 CTTTTCGTTT TCAAAGGATC AAGAACACCT TTAAAGGTG TTTTGGCATA
 216251 TTTCTCATAA TCGTTTAGAG AACATTAGTG TTCGAGATTT TGTCTGTAGT
 216301 TATAAGTATC AGGAGTATTT TGGGCCTAAG CCTTTCCTA GATGGAGAGT
 216351 TCAGATTTTA GAATGGCGGG GTTATGTTAA AAAAGAACAA GATTATTATC
 216401 GACTCACAAA AAAAGGAAGA AGTGAGGCCT TAAGATTAGT TCGTGCTCAC
 216451 AGATTATGGG AATCGTATCT TGTGAATTCT TTAGATTTTA GCAAGGAAAG
 216501 TGTTCATGAG TTGGCTGAGG AAATAGAGCA TGTTCCTACT GAAGAATTGG
 216551 ATCATACCTT GACAGAGATT CTCAATGATC CTTGTTATGA TCCTCATCGA
 216601 CAAATTATCC CAAATAAAAA AAAGGAAGTC TAATGGCTTT GGGACCTTCT
 216651 CCTTATTATG GAGTATCTTT TTTCCAATTT TTTTCAGTAT TTTTTTCGAG
 216701 ACTGTTTTCT GGAAGTCTTT TCACGGGTTC TCTCTATATT GATGATATTC
 216751 AGATTATAGT ATTCCTTGCT ATTCCTGTT CAGGTGCTTT TGCAGGAAGT
 216801 TTTTLAGTCT TGCGAAAGAT GGCTATGTAT GCGAATGCTG TCTCTCATAC
 216851 TGTCTTTTTT GGTTTGGTCT GTGTTTGTTC GTTTACGCAT CAACTGACGA
 216901 CCCTCTCTTT GGGTACCTTG ACTCTTGCAG CAATGGCAAC AGCTATGCTG
 216951 ACAGGGTTTC TTATTTACTT TATTCGTAAT ACTTTTAAAG TTTCAGAAGA
 217001 GAGCAGCACC GCTCTAGTCT TTTCTTTATT ATTCTCTCTG AGCCTTGTTT
 217051 TGTTAGTCTT TATGACAAAG AATGCTCATA TAGGAACGGA GCTTGTGTTA
 217101 GGAAACGCAG ATTCTTTAAC GAAAGAGGAT ATTTTCCCTG TCACTATTGT
 217151 GATTTTGGCT AATGCTGTAA TTAATATTTT TCGGTTCCGT AGCTTAGTTT
 217201 GTTCTTCTTT CGATTCTGTA TTTGCCTCTT CTTTAGGAAT TCCTATTCGG
 217251 TTGGTTGATT ATTTGATTAT TTTTCAACTT TCTGCATGTC TTGTAGGAGC

217301 TTTTAAGGCT GTAGGTGTAT TAATGGCACT TGCTTTTCTG ATCATTCCAT
 217351 CGCTTATTGC TAAGGTTATT GCAAAATCGA TAAGGAGTCT TATGGCTTGG
 217401 TCGTTAGTTT TTAGTATTGG TACAGCATTT TTAGCTCCTG CATCTTCGAG
 217451 AGCAATTCTT AGTGCTTATG ATTTGGGGTT ATCGACTTCG GGAATCTCTG
 217501 TAGTGTTCCT GACGATGATG TACATCGTGG TTAAATTTAT AAGCTATTTT
 217551 CGAGGCTATT TTTCTAAAAA TTTTGAAAAA ATAAGTGAGA AAAGTTCTCA
 217601 ATATTAGCAG TGATTTAAAA GAATGAAATT TTAGAGTTTA GTCCACTTTA
 217651 GTTAGAAATA GTATCAATAG AGAATTGACA ATTCCTCGAC TTGCGGAGTA
 217701 TGATTCTCTT TTTCACCTAG TTAAAGGTAG CATGCTTGAA ACATTTAGCC
 217751 GTTCTTGGGT CAACAGGTAG TATTGGCCGT CAAACATTAG AGATTGTGCG
 217801 GCGCTATCCT TCAGAATTTA AAATTATTTT TATGGCTTCT TATGGAAATA
 217851 ATCTAAGGTT ATTTTTTCAG CAACTAGAGG AGTTTGCTCC GTTAGCCGCA
 217901 GCGGTCTATA ACGAAGAGGT TTATAACGAG GCCTGTCAGC GATTCCCCCA
 217951 TATGCAATTT TTCCTAGGCC AGGAGGGTTT AACCCAACTT TGTATCATGG
 218001 ATACAGTCAC TACTGTCGTT GCTGCTTCTT CAGGAATCGA GCGCTACCC
 218051 GCGATTCTAG AGTCGATGAA AAAAGGAAAA GCACTAGCTT TAGCAAACAA
 218101 AGAAATTTTA GTTTGTGCTG GCGAATTGGT TTCTAAGACT GCAAAGGAAA
 218151 ATGGTATAAA AGTTCTTCCT ATTGATAGCG AGCATAATGC TTTGTATCAA
 218201 TGTTTAGAAG GCAGGACGAT TGAGGGAATC AAGAAACTGA TTCTTACAGC
 218251 TTCTGGAGGG CCTCTGCTCA ACAAGTCTTT AGAAGAGCTT TCTTGTGTAA
 218301 CAAAACAAGA TGTTTTGAAC CATCCTATAT GGAATATGGG TTCAAAGTG
 218351 ACTGTGGACT CATCCACATT GGTCAATAAG GGAATCGAAA TTATCGAGGC
 218401 GTATTGGCTG TTTGGTTTAG AAAATGTTGA AATCCTGGCT GTAATTCATC
 218451 CTCAGAGCTT AATCCATGGT ATGGTAGAGT TTTTAGATGG GAGTGTGATT
 218501 TCTATCATGA ATCCGCCTGA TATGCTCTTC CCAATACAAT ACGCTTTAAC
 218551 AGCTCCAGAG CGTTTTGCAT CTCCTAGGGA TGGTATGGAT TTTTCGAAGA
 218601 AACAACTTT AGAATTTTTT CCGGTAGATG AGGAGCGATT TCCTAGTATC

218651 CGTTTAGCAC AACAGGTATT AGAGAAACAG GGGTCTTCTG GAAGCTTTTT
 218701 TAATGCAGCC AATGAAGTAT TAGTGCGGAG GTTCCTTTGC GAAGAGATTT
 218751 CTTGGTGTGA CATTTTACGC AAATTAACGA CTCTTATGGA ATGTCATAAG
 218801 GTTTATGCCT GCCACTCTTT AGAAGATATT TTAGAAGTAG ATGGTGAGGC
 218851 TAGAGCTCTT GCTCAAGAAA TATAATCGAG TAGGTATATG ACAATAATCT
 218901 ATTTTATTCT AGCAGCCCTA GCTTTAGGGA TTTTAGTGTT AATTCATGAA
 218951 CTTGGTCATC TGGTAGTAGC AAAAGCTGTA GGAATGGCTG TAGAGAGTTT
 219001 TAGCATAGGC TTTGGTCCTG CTTTATTTAA AAAGCGTATA GGCGGCATAG
 219051 AATATCGCAT TGGATGCATT CCTTTTGGAG GCTATGTTCTG TATCAGAGGT
 219101 ATGGAACGTA CCAAAGAAAA AGGGGAGAAG GGAAGATAG ACTCTGTCTA
 219151 TGATATTCCT CAGGGATTTT TTAGTAAGTC TCCTTGGAAA CGCATTCTGG
 219201 TTCTTGTTGC TGGTCCTCTT GCCAATATTT TATTAGCTGT CTTGGCTTTC
 219251 AGCATTCTTT ACATGAATGG GGAAGAAGT AAAAATTATA GCGACTGTTC
 219301 TAAAGTGGTA GGTGGGTCC ATCCTGTTTT ACAGGCAGAA GGATTGCTCC
 219351 CTGGAGACGA GATTCTTACG TGTAATGGTA AGCCTTATGT GGGAGATAAG
 219401 GACATGCTAA CAACCTCTTT ATTAGAGGGG CATCTCAATC TAGAAATCAA
 219451 ACGTCCTGGC TATTTGACAG TTCCTAGCAA AGAGTTCGCT ATTGATGTTG
 219501 AGTTTGATCC CACAAAATTC GGGGTTCCT GTTCTGGAGC GAGTTATCTT
 219551 TTGTATAGCA ACCAGGTGCC CCTAACGAAG AACTCTCCTA TGGAGAATTC
 219601 AGAGCTACGT CCGAATGATC GTTTCGTTTG GATGGATGGC ACACTTCTTT
 219651 TCTCAATGGC TCAGATATCT CAGATACTCA ATGAGTCTTA TGCTTTTGTG
 219701 AAAGTAGCAC GGAATGACAA AATCTTCTTT TCTCGTCAAC CTAGGTATT
 219751 GGCTTCCGTT TTACATTACA CTCCCTACCT TCGTAATGAG CTTATAGATA
 219801 CGCAGTATGA GGCTGGACTT AAAGGCAAGT GGTCTTCGTT ATATACATTG
 219851 CCTTATGTAA TCAATAGTTA TGGATACATA GAAGGTGAAC TTACTGCTAT
 219901 AGATCCAGAG TCTCCTTTGC CACAACCTCA AGAGAGGCTA CAGCTTGGGG
 219951 ATCGCATTCT AGCTATTGAT GGAACCTCTG TTTCTGGAAG TGTAGATATT

220001 TTACGTCTTG TTCAGAACCA TCGGGTCTCT ATTATTGTTC AGCAGATGAG
 220051 TCCGCAAGAA CTTGAAGAGG TGAATTCTCG AGATGCTGAT AAGCGGTTTA
 220101 TCGCCTCTTA TCATTCCGAA GATCTGTTAC AAATTTTGAA CCATTTAGGA
 220151 GAGTCTCACC CAGTAGAAGT CGCGGGTCCT TATCGTCTTC TTGACCCTGT
 220201 TCAGCCTCGT CCTTGGATTG ATGTTTATTC TTCGGAGAGT TTGGATAAAC
 220251 AGTTGGAAGT AGCTAAGAAG ATTAAGAACA AGGATAAACA AAGATACTAT
 220301 TTGGAGCGTC TTGATGCTGA GAAGCAAAAA CCATCTTTAG GGATTTCTTT
 220351 GAAAGATCTT AAGGTGAGGT ATAATCCTTC ACCTGTGGTT ATGTTATCAA
 220401 ATATTACTAA GGAAAGTTTG ATCACCTTGA AAGCTTTAGT TACTGGACAT
 220451 CTGAGTCCAC AATGGCTTTC AGGACCTGTG GGTATTGTGC AGGTTTTACA
 220501 TACAGGATGG TCGGTAGGGT TTTCTGAAGT GCTCTTTTGG ATCGGTCTAA
 220551 TTAGTATGAA TTTGGCTGTC TTGAATTGTC TTCCTATTCC TGTTTTGGAT
 220601 GGAGGTATA TCCTTCTATG TTTGTGGGAG ATAGTCAAAA GAAGACGTTT
 220651 GAATATGAAG ATTGTGGAAG GGATTTTGGT TCCGTTCACT TTTTTATTGA
 220701 TCATCTTCTT TATTTTTCTA ACTTTTCAGG ATTTATTTTCG TTTTTTTGGT
 220751 TAAGGCTCCA TTTTTTTGGA AATCCGTAAA GGTTGTAAGG ACGAGAACCA
 220801 AAGCAGTGTA TAACGTAGAT GTCTTGTTA GGGTAAGGCT TAAGTCTTGA
 220851 GTTCTTTTGT TTGTCTTCCA TAACAGTTTG TTCTAAAGCT GCTTCGGGAA
 220901 CGTGCTCGAT TATCGGGGTT GTATTTTCCTT CTTTTAAAAA CTCTTTCATA
 220951 AGATAGAGCG ACTTCCGATT AGTAATGATG TAGGTATGAC TCGTGTGGAT
 221001 CACGTAAGCT TTATGTGGAG TATCTAGCTT TGTCTCTATA GTGCATACCT
 221051 TACCATCGTT TTCATAGGGC AGGCGGAAAG GAAGGAATTT GCTATGTCTG
 221101 TTCCCACTAA GAATGAGAAC GTCTAAAGAC GAAGGGAGTT TCCCGACATT
 221151 TAACATCTTT GTGGGGCAGT AGGTAAGAAG CTGTCGTCCT AATTTTCCTC
 221201 CAAATACGAA CTGTACGAAT TTCACACAGC GGTAGCGTCT AGCTAGTGTA
 221251 GACCCTGCGT TCGGAGGAGC CATGAGAATA GCTTTTCTTT TTTGGCTTCT
 221301 TCGGGGCAAT CAGGGTGAGC GAGTGCTACA CGAACAATGA CTCCTCCAAT

221351 AGAATGAGTT ACGAAGTTTA TAGGGACTCC AGGCTTAAGT TCAGCTATTT
 221401 TTTTCAGCAA GCGATTGAGA TGTTTCAGCAT GCTTTTCTAG AGTAAACTTG
 221451 CGCGTCTCAT AATTCCAAAT AAAGACATCG TAATGTTCTT TTTCTAGAAC
 221501 GCGAGCAATA GGTTTTAAAG ATGTATAAGA TCTTAAAAAC GCATGCACGC
 221551 AGACCACGGA TTCTTTTTGT TTTGAGGTTT CCTTTAATCC CCCAATTCCA
 221601 GATGGAAGGG TTTGGATTAC CGAGGTTTCC GAGAATAAAG GATTACCCAC
 221651 AGCTAAAAAG AGTATAGTTA ATAAAAATTT CTTCATGTGT CACCGATATA
 221701 ATTATTTTAA ATAAGTTTTT TATTAAAAAA GAATATGTTT ATTAAACTT
 221751 ATTTATAAAA ATAACTAATT CGCTTACTTT TAACGCAAAA GAAAAATGAT
 221801 GTTATTGTTT TAGATAATTT GCTCAGAAAC TTGAGCGTTC TCGATACTTA
 221851 AAACAAGGCT TGTTTTTGA AGTTCCCAT GCATATGGGT GGAAGTAATC
 221901 AGAGTCTGAC CCAGAGTTGG GGCAGGGTCA AGGAGTTGAC CGACACGTC
 221951 ATTATCTAAT CCAGCATGGA TATCATCTAG ACAGACTAGA GGGGAGACGT
 222001 GATGAGATTG CTTTAGATAT AGGCACTCAG CAAGCCTTAA GATTGCCAAA
 222051 AGACTGTGTT TCTGCCCTTC ACTAGAGAAT TGAGACACAG GCATTTGGTT
 222101 CATAGTGAGT AGAAAGTCTT CGCGATGAGG GCCAACGGAA GTGCTTCCTC
 222151 ATTCGAGATC TCTAGGAAGT GATATAGAGA GCTGTTTATG AAATTCTTCG
 222201 GCAACAGCAG TTTCAGAAAT ATCAGAATTT TTAATTAAGG AACTTTTAAA
 222251 TTTTAAGGCC AATTGTTCTT TTAGGTTGTT AGACCAAAGT TCTTTGGATA
 222301 AATCTGAAAG TTTCTGACTA CAGAGAAACC GTTGGATGGA TAGGTAGGTG
 222351 CCGTGTTTGA CCAACTGTTC ATCCCAGAGG CCACGGTTGA GGTTTGCTTG
 222401 CTTTTTAAGA GAGCATTTCT CTGCTGAAGA GCGCGATGAT AGTACGATAA
 222451 GCAGAGGGTA TAGTGGTTAT CGCATTGAGA TAAAAGCAGA TTTAGGAAAA
 222501 GGCGACGATC CGCAGGAGCT CCTGAAATTA GAAGGCGGTC TTTTGAAGAG
 222551 AAAAGCACAA TAGGTACTTT CCCTATCAGC TGCATAAGG TTTTATAGG
 222601 AAGTTGTTA TAGCAGATTT TTTTTCCTTG CTTGTCTGTA TAGATGGAGA
 222651 GAGCTTGGGG AAGGTGGTCT TTCTCAAAC GTGTTTCTAA GAAGAAATGG

222701 GAAGATCCGA AGGTGATGGT ATCTGTGAGA TGTTGCGTGC GAAAAGACCT
222751 TCCCAAGGAC AAAACATAAA GCGCTTCTAG GAGGTTTGTT TTCCTTGGG
222801 CATAATTGAG TTTAGGAGCC AGTGAGATTT CTAAATCACT GTGGTTACGA
222851 AAATTTTTTA GCTTCAGAGA GCAGATTTTC ATAAACATCG GCAGGGTAGT
222901 GACCTAAAGG ATTCTCCTTA GGGAGTTTAT TAATCATCAT GTAGCCTCAT
222951 AGGCATGATG ACAAATAATC CTGAGGCAGA ATCGGTAATG ATTCCAGGAT
223001 TATAGGAATC CGAGATCCCT AAGCTGACTA ATTCATCCTT ACTATGCTTC
223051 AGGATATCTA AAAAGAAAAA GGGATTAAAG GCAATTTCTA GGAGTTCGCC
223101 AGAATAATTT ACAGCCATGC TTACCTTTCC TTCACCCACC TTAGTACAGT
223151 TGGCTGTTAG AGTGAGCTCT CCGGGTAAGA AAGAAACTT CACGGAGTGA
223201 GAGGACTCAT TTGTAAATAA AGCCACTTGT TTGAGCAGAG TAATTAGTTC
223251 TTCGCGATGC AGATCGAGTT TTACGTTGCT TTCTGTAGAT ATGACGGGGG
223301 AGAAATCTGG AAATTCCTCA GAAAGAAGTT TTGTGATCAG GAGAGTATTG
223351 TCACATTCAA CCGCAATCTT ATCTTGATCC AAGAAGATCG CAGCTTCACC
223401 TTCATCGGAG CACATCTTTA TAATTTCTTC TACTGCTTTG ATAGGAATAA
223451 TATATTCCCC AGAAAACTT TTATCTAAAG TAACTTCAGC ATCTATTTTT
223501 GCTAAACGCT TTCCGTCAGT CCCTACGATG GTAGCCACGC CATTGGCGAT
223551 AGCAAGCAGG ACTCCAGTAA GAACATAGCG GCTTTCTTCT CTAGATACAG
223601 CGAATGAAGT TCTCTGTAGC ATGGTTTTTA GCTGCTCTGC AGGCAAGGAA
223651 AAACGCAAAG CATTTTGTAT ATCAGGGAGC ATGGGGAAGT CTTCTTTTTC
223701 CATGCTGAGT AGGCGAAAGC ATGAAGATCC CGAGGTGATT TGTGCCATTT
223751 CCCCTGCTGA AGAGGAAATT TCTAAATTTG CCTCTGTTAA TTCTTTTACT
223801 AATTGAAAAA ATCTCTTGGA GGGAATGGAA ATAGCGCCTT TCTCATAGAC
223851 TTTAGCTTTG GTGACGCAAC GTGTGCTCAC TGTCAGATCC GTAGCAGTGA
223901 AAATAATTC ATCATTATAA GTTCAATCA AAACATGGGT GAGTACTGGA
223951 ATAGGTGTGT TTTGAGGGAC GACACTTTGA ATTTTTTTGA TAAGGTTTCC
224001 TAGCTCATTT CGGGATACAA CGAATTCAT ATTTTCCTAT AACCTGAGTC

224051 TAACTTTATA TGAAGGGCTA CGCCTCTCTA GGAGAGACGG TAGAGGTACG
 224101 ATCGGGATAT TACTACAAAG TGATTTTCTG GTCTTCTAGA TATTTTACTA
 224151 ATAAAAAAGG CTTTAGATTG AGGAAATCTT CCCTGGAAAT CAAGGAAAGA
 224201 GGATTCTAAT CATTGTCTTA AGAAGGAAAA ATTGCTTTTT GTTATACTGG
 224251 TTTTGTCCC CAATTATGGG GGAGGATCTT ATGGCACAAA AAGAAATTGT
 224301 TTCTAATCGC AAGGCTCTGC GTAACATGA AGTTATAGAG ACTTTAGAAG
 224351 CAGGCATCGT TTTGACTGGG ACTGAGATTA AGTCGTTGCG CGATCATGGG
 224401 GGAAACCTCG GTGATGCTTA TGTCAATTGTT TCTAAAGGTG AGGGGTGGTT
 224451 ATTAAACGCG AGTATTGCTC CCTATCGGTT TGGAAATATC TATAACCATG
 224501 AGGAGCGTCG TAAACGTAAA CTCCTTCTTC ATAGATATGA ACTTCGTAAG
 224551 TTAGAGGGTA AGATTGCTCA AAAGGGCATG ACTTTGATTC CTCTGGGAAT
 224601 GTTTCTGAGT CGCGGCTATG TTAAGGTACG TTTGGGTGTG TGTCGTGGGA
 224651 AAAAAGCTTA TGATAAGCGT CGTACGATCA TAGAAAGAGA AAAGGAACGT
 224701 GAAGTTGCCG CTGCTATGAA GAGGCGCCAT CATTGATATA GGTTAGGATA
 224751 TGGTGTTCTT CAGCCCACTG TTTTGCTTCT ATTTTAGAAT CAAAAGTCAT
 224801 GAGGACTGTG GCAATAGCGT CGGCGTATGC GCAGCTCGGA TGGACTACTG
 224851 AAACACTTTG GATAGGATAG GAGCTTAGCT CTAGGGGTTT CCCTGTACGA
 224901 GTATCAAGAA TATGGGTGTA AATTTTTCCT TCAACACACC ATTTTGAAT
 224951 ATGATTTCCA CTTGTTGCAA TTGCCATATC ATCGATATCT AAGATCGTAC
 225001 CTGCTGCTTC AGAAAAATA CGCCAAGGTC TTCCCGAGGG ATGATGCCCT
 225051 GACGTTTTGA TCTCTCCTCC CCACTCTACA TAGTTGTTG GACAAAAGGT
 225101 ATTGCAAATT TCATTTAGAC AATCTACGGC ATAACCTTTG ACAACACCAC
 225151 AGAGGTCGAT TTGAACATGA GGATTCTTTT TGATTAGAGT TTTTGTGTTT
 225201 GACTGAAACT CCAAGTGTTG CCAGCCCATG TCTTTATAAT GTTGTTCCTA
 225251 AACGTCTTTA GGGGGGAGGG TTTGACTTTT GAGATGTAGA AGCCATAGGG
 225301 TTTTAAAGG TCCTACAGTA GGGTCAAAC GTCCTTCTGA AAGTTGTAA
 225351 AGTGTATCTA CCTGATCTAG AACTCGGAA AGTTCTACAG ATAAAGTTAT

225401 GGGGACATCT GCTGGAGCTC GGTGATTAT CGAGAGTTCA GAATAGGGAT
 225451 TCCAGTTGTT ATAAATCGAG TCGATCTTAT GAAAGCATCT ATCAATTTGT
 225501 TGGGATAAAG ATGCTTTTTC TTTTGCGGAT AAAGAGGTTT CCAGAACAAT
 225551 GCGATAGAAG ATTGTCATCT GCTCTCCTTC GATTGTTGTC GTTTTTTGAG
 225601 AGCATGAACA GAGTCCAAGA CATAAAAGAA CTAAGAAAAA TTTTGGTAAC
 225651 ATCGCCATGT ATTGCTCATG AGCATAGCGA CAGTCATGGG ACCAACGCCT
 225701 CCAGGAACTG GAGTGATTGC TCGCATTTT GTCACAACGT TATTAAAATC
 225751 TACATCCCCA AGAAGAGTAT AGCCTTTCGC ATTGTCTGCA GGGACTCTTG
 225801 TTGTTCTTAC ATCTACGATC ACAGCATGTG GGGCTACCAT AGTTTCCTTT
 225851 ATAAAAAGCG GTGCTCCTAG AGCAGCAATA ATGATATCAG CTGTCTTTAA
 225901 GATTTCTGGG AGGTTTTCCG ACTGGCTATG AAGAACTGTG ACTGTACAGT
 225951 TAGTTTGAGG ATGCTTTTGC ATCATGAGGG CCGCTAAGGG TTTCCCCACG
 226001 ATGTTGCTTC TCCCTACAAT AGCGGCATGG CGGCCTCGAA GAGGAATTTT
 226051 ATAATAGTTC AGGAGTTCAA TAATTCCTGC AGGAGTGCAG GGTAGAAGTC
 226101 CATCAAAATT TCCAAGGAGC AACTTTCCCA TGTTTACAGG GTGAAGCCCC
 226151 TCCACATCTT TGTCTGGGGA GATCGCTTGG AGAATCACTT CGCTGTCCAA
 226201 GTGTTTGGGC AAGGGAAGTT GCACGAGGAT GCCGTGGATG CTAGGATCTT
 226251 GATTCAATCG TTCTATGAGC TTAAGGACTG AGGAGAGGGT AGAGTCAGAG
 226301 GGTAACCTGT GCGCTTTGGA GATAATTCCG ATTTCTGTAG CTTTTTTGAC
 226351 TTTCATGCCA ACGTACACCT CAGATGCGGG GTCATTGCCA ATCAGGACCA
 226401 CAGCAAGCCC CGGAGAGGTA GGACTTTGTG AGATTTCTTC TTTGAGTCTC
 226451 TGAAGGATTT TTTCAGCTGC AGGAATCCCT CTCAGTAACA TACCAATCTC
 226501 CTAAAAAAG ATATCATTTT ATGGATTCCCT GAGTTTATTG CGCAAGGAAG
 226551 TATATAGATT ACAGAAGTGA AGATGAGATT ATCTAAGAAA GAACATTTTA
 226601 GAGGTATCAT TAGCACATCT ATTGACCAAC GGACTTCATG TCCGAACAAG
 226651 GCGCACATAG GAATGGTTAG GAGTAAAAAG ACTCCATAGA AAATCATATT
 226701 CACAATGACC AAGGCGGCTT TGCTCTCTTT TGAAAAAATG AGGTGTGCCT

226751 TATGGGTGAT CAGAGCAGTT AGAGGGTAGA GCAGTAGAAA TACACCGCGG
 226801 CTTCTTAAGG CAAGATCACT TAAGACACCA CAGCCTAGAG CGAGTACCAG
 226851 TCCCTTATCT TTTGAGAAAC AATAGAAACA AAGAACAATA TAAGGGGAAA
 226901 AAAGGATCAG ATGTGCTTTA GGGAAGAGCG TTGAAGGGT GAACATGGAT
 226951 AAAATACACA AAATAAGCA TTTGCGCGGT GTTATGTGGA CAGACATGGT
 227001 TTACGCTTTG AATTATGATT ATCGATTTC AAGGTTTTG CAGACTTGTA
 227051 TATAGCAAAG CATAGGATTC TTAGCTTGAT TTTTATGTT TATCTAACT
 227101 GCTTTTAAAC AATTATTTAT ATAATATTTT ATTTTATAGT TAAATTAGT
 227151 TTATTCTTAA TAAGTTTTTT AATTAAAATT AGTTTATTAA GTTGATTTTT
 227201 TTGTTTTTGT ATTTTTTATG GGGAAGCCTA AGAAGAGCAG AACGGATAGG
 227251 GCTTTGGCTC AGGAGATTCA AAAGAAATCA ACGGAAGTGT TGAAGAAGCC
 227301 TGC GCGGATA AAAGCTAAAA ATCGTCGTAA ATTTCTTATT GCTAAGGAAC
 227351 AGAAACTCT TAAACACCGT GCTCAAGAAT ACGATCAGTT AGTTCGCTCT
 227401 CTCTTAGATT CTCAGAAGAA GGACACCGAT AAAGTTTGA TTTTCAATTA
 227451 TGAGAATGGG TTTGTTTTTA CTGACAAGGA CCATTTTAGT AAGTACTCTA
 227501 TCCGTCTTTA GCGGGGGTTA AATTAAAAAA AAATCGCAGA ACGGGACTAT
 227551 CCCGTGATGG AACGACAGGA AGCCTGGGTA ATTAATTTAG AGTTAGGAAC
 227601 TGTCTCGGTA TTAAGATCTC CAAAATCTTA ATACCTAGAC AGTCGGATCC
 227651 GAGTGATCCC TCTAGCTCGA CAGCGATGCA CTTTGTTGCA TCTACTCGCT
 227701 GTTCTCTGTC CTCCCATAAG TTTTTTTACG CAAGGTGTTT CACCTTGCCT
 227751 TTTTTTTTGT TTTTATAGTT TTTAAAATTT GTTTGTAAGT TCTTTTTTTT
 227801 AAATTAATTA TTTTTTGTTT TAAAATTATA AGTAGTTATA AGTTTTATTT
 227851 TTTAAATTCA GAGAATCATT ATGGCAGTTT CAGGTGGCGG AGGGGTTCAG
 227901 CCTTCTTCGG ATCCAGGAAA GTGGAATCCT GCTCTGCAAG GAGAGCAGGC
 227951 AGAAGGCCCG TCTCCGCTAA AAGAATCTAT ATTTTCTGAA ACCAAGCAGG
 228001 CCTCCTCTGC TGCGAAGCAG GAAAGCTTAG TGCGTTCAGG ATCTACAGGA
 228051 ATGTATGCAA CAGAATCTCA GATAAATAAG GCTAAGTATC GTAAAGCTCA

228101 AGATCGATCA TCAACCTCTC CAAAATCCAA ATTGAAAGGT ACATTTTCTA
 228151 AAATGCGCGC TAGTGTGCAA GGATTCATGT CAGGATTCGG ATCTCGGGCT
 228201 TCGAGAGTGT CAGCAAAGCG TGCTTCCGAT AGTGGTGAGG GAACATCCTT
 228251 ATTGCCGACA GAGATGGATG TTGCTCTAAA GAAGGGAAAC CGTATTTTAC
 228301 CTGAAATGCA GGGATTTTTC TTAGATGCTT CGGGTATGGG AGGGAGTTCC
 228351 TCTGATATTT CTCAGCTTTC TTTAGAGGCT TTGAAATCTT CAGCATTTTC
 228401 AGGTGCCAGG AGTTTAAGTT TAAGCTCTTC AGAATCTAGT TCCGTGGCTT
 228451 CGTTTGGATC TTTCCAAAAG GCCATAGAGC CTATGAGTGA GGAGAAGGTA
 228501 AATGCTTGGA CAGTGGCTCG TTTAGGAGGG GAGATGGTCA GCTCTCTTCT
 228551 CGATCCCAAT GTTGAGACCT CATCATTAGT GCGCAGGGCA ATGGCAACAG
 228601 GCAACGAAGG CATGATAGAT CTTTCTGATT TAGGACAGGA AGAGGTCAGT
 228651 ACAGCCATGA CATCTCCCAG AGCAGTAGAA GGAAAAGTAA AGGTATCTTC
 228701 TTCTGATTCT CCAGAAGCGA ATCCAACAGG AATTCCAAAT TCTAATACTT
 228751 TAGAAAGGGC GGAAAAGGAA GCAGAGAAAC AAGAAAGTCG AGAGCAGTTG
 228801 AGTGAGGATC AGATGATGCT TGCACGTGCT ATGGCTGGGC TTCTTACAGG
 228851 GGCAGCGCCT CAAGAGGTAT TGAGTAATTC TGTTTGGTCT GGTCCCTTCTA
 228901 CAGTATTTCC TCCTCCCAAG TTTTCAGGAA CTTTACCCAC CCAGAGATCG
 228951 GGAGATAAAT CAAAGCATAA ATCTCCAGGA ATAGAGAAGA GTACGAACCA
 229001 TACGAACTTT TCTCCTCTTC GGAAGGTAC TGTGAAGAGT GCTGAGGTTA
 229051 AAAGTTTGCC TCATCCAGAA AGTATGTATC GTTTTCCTAA AGATAGCATC
 229101 GTTTCCAGGG AGGAACCTGA AGCCGTTGTT AAAGAATCTA CGGCATTCAA
 229151 AAATCCAGAG AATAGCAGTC AAAACTTTCT CCCTATTGCT GTGGAGAGTG
 229201 TTTTCCCTAA GGAAAGTGGT ACGGGAGGGG CTTTAGGAAG TGATGCTGTG
 229251 AGTTCCTCAT ATCATTTCTT TGCGCAACGT GGAGTGCTTT TACTCGCTCC
 229301 TCTACCTCGT GCTACTGATG ACTATAAAGA GAAGCTCGAA GCTCATAAAG
 229351 GTCCTGGAGG TCCTCCAGAT CCTTTGATTT ATCAGTATCG AAATGTTGCT
 229401 GTTGAGCCGC CAATTGTTCT CCGTTCTCCC CAGCCGTTTT CAGGATCTTC

229451 ACGTCTATCG GTTCAAGGAA AGCCTGAAGC TGCTTCAGTT CATGACGATG
 229501 GTGGGGGGGG AAATAGTGGT GGTTTTAGCG GAGATCAAAG AAGAGGATCT
 229551 TCGGGCCAGA AAGCTTCCCG TCAGGAAAAG AAGGGAAAAA AATTATCTAC
 229601 GGATATTTAG GGTTTTAAGT CGGTTTGATG TGATGCGTAT AATTCGTTTT
 229651 GATCCTTATG GTGCGCTATC TGCACAAAGC ATAGCTAAAG ATTCCCGTCA
 229701 AAACCTCCTT TTAGTAGAAA AAATTTCTGA GGAAATTGCT ACGAATGAAG
 229751 CGATTGCGCT TGCCTTGCTA GCTATTGGAG ATCGCGAACA AGAGGAGAAG
 229801 AAACAGAGGC ATCGTTATAA GCTACTCGGA CAAAAGCAAG CCAAGGTCTT
 229851 GCTTTCTCAG TTGCGTCATG TGCATTTAGA TTTTAAAAAA CTATATTGCG
 229901 ATAGTAAGAA AAAAGAAGAT CAGGAAAAAG ACGAAAAAAA CAAACAGAAG
 229951 CGATCTATTA AAGTTACAAA GAAAAAAAAG GGCATCTCTT TAGGGGCTGC
 230001 CGCTTCTCAG GCAATTGCAG CAGCAGCAGA AGCTTGGGTA ATTGCTAGAA
 230051 ATAAAGGAGT CTTAGAAACT GCCTCCACTC TTTTTTATCA AAAGGATGAA
 230101 GAGGCCTAGA CATACTTGAA TCACGAGCTA GGCCTTCTTC TGTGACTATT
 230151 TTAGATCATC ACGCTGCTTC TTGCTCTTCT ACTGGAAGAC TTTTTTCTAT
 230201 AATTTCTTTT TCTTCATCGT CTA CTGTAGG ATGTTCTGAA TGCTTAGCTA
 230251 GATCTTTCCA AATCATTCGA AGTTTCTGAT TTTGTTGTTT GGTGAGGTTT
 230301 TCCAGTATGA TCAAGCTTTC ATCATTTAAG GAGAATCTTC CTTTAGACCA
 230351 ATTTATAGAT CCTGCAAGTA GAGTTTATT ATCTATAACT GCAAACCTAT
 230401 GGTGAAGAGT ACAGGGTGCG GTATTTATAG AAACAAAGTC TTTATTGATA
 230451 TTTAATTGTC GTAATTGCTT AAAAGTAAGT TTGCTATGAC TTCTATCAAT
 230501 GATAATATCT ACATGGATTC CTCGTTGTTT TGCTTGATGT AAGGCTTGAA
 230551 TAATCTCCGA GTGGGTCAGA GCAAACATAG CAACTTGGAT GGTTTTCTGA
 230601 GCTGTCTGGA TTTTTCGAG TACAGCTTGT ATTGCAATTT TACGATCTTG
 230651 AGGAAGAACA AAATACTTTC CTGTTTGATC CTTTATAGAA AAGTCTCCAG
 230701 AGGTATTTGT GATAATGAGA TCACAGAGCT CCGAGCTATG CATTCCTAGA
 230751 ATGAGATTAT TATCTAAACG TAGAGAAAGA TTGGTGTAGT TCGCAGATCC

230801 TAGCCAAGCA TCTTTCTTAT CTATGGAAAG AGCTTTTTGA TGCATCAGTT
230851 TACGCCCTGC TGGAGGTTGC TCGACTAAAG TTACATTGCT GGCTTGCTTT
230901 AAGATTTGGG GAATTTTAAA TTTTGTATAG TAGATCGTAA CTTTGTTTTT
230951 TGCTTGAGCT TGTCGAGTTA AACTCTGTTG GATCTTGGGT TCTGAGAGGT
231001 TATAAATACG TAGGAAGATC TCTTCATCAG CGTGTCTTAT AGCATCGCAT
231051 AGAATTTTAC GCATGTCCTC ATTGCATTGA TTTGAGTAGA TGATAGCTTC
231101 TTCAGACTTT AAAAAAGTCT TAAAAGTGTC ACCACGAGGA GCTCTTGCAA
231151 AAATTCCTAC TAAAATCAAC GTGCTAATAA TAACACAGAT TTTTAATTTA
231201 TCTTTTTGTC TTTTATTCAT TTTTACTCT TAAAATAGCT TTTTGTTTT
231251 TATATATTTT TAAAAATAAA AATAACTAAT TTAATTAAAC GTTTAATAGA
231301 ATAAGATTCT TATTAAATAT AAAAAGTTTC TTTATTGGGG AAATATGCGT
231351 GCCATCGTTC GGAAACCTTT TGTTTTGTAG ATGGGTCGAC CTCTACTTCT
231401 TTAGGATAGG AAGGCTTCAT CAGGGCATCG ATAACGAAGG GAAAGTTGTA
231451 ATTAGGACGG TGAGTAGCAA AATGGCTGTG GAGCGCGTGA AGATCATTTG
231501 CTGGGGCACA TCGTGTGAAG GTCCTCCAGA GAAAATCTTT TTCACTTTGA
231551 ATGGTTTCTC TCAGATTATC GGCAAGGATA ATCAGAGGCC ATGATTTTAG
231601 ATCTGGATGG TGAAGGAGAG ATTTAATACA TCGGTCCTCG AGGGATGTTT
231651 CCAACACTAG GCAACCACGA CAAAAGGGAG CGATGTCTTG AACTCCATGG
231701 ATTTTTCTC CTGATATCC ATGGGAAGG TCTCGGATGG CTTTCTCTAT
231751 TCCCATGAAG ATTCCCTTGG AGCCCTTATT TAAGCTTGGT CCTGTATAGT
231801 CTAACGTATC GTTTCAGTT TCTGAGAAAA TAATAAGATC TCGGTCTGGC
231851 TGTAGACGCT CTAAAATGGT TTCTAGAACC ACGGAGAACC TGTCGAGAGG
231901 CACCTCTTGG TCTGTGACCA TTAGGAATTT CGTTAGGGAA AGTTGGCCCT
231951 CTCCAAGAAT TCTAAGAGCT GTGGTTAGAG ATTCTCTCCA ATAGCGTTCT
232001 TTAACGACAG CGGCAGTCAG TGCATGAAAC CCTGATTCTC CGTAACTTTT
232051 AAGTCTACGC ACACCAGGCA TAACTAACGG AAATAAAGGG GAGAGGTATT
232101 CTTGGAGTTT GTTCCCTATA TAAAAATCTT CTTGGTAGGG TTTGCCGACT

232151 ACTGTAGCAG GATAGATTGC ATCTTTTCTG TGATAGATTT TATGACAGTG
 232201 GAATTCAGGG AAGTCATGTT GGAGACTGTA GTATCCGAAA TGATCGCCAA
 232251 AAGGACCTTC AGGACGACGT TTCCCGGCCG GAGATTCTCC GACCAGGATG
 232301 AATTCCGCAT CGTAGAGTAG AGGGTGGGGA TGGTCGTTTG TTTTTTTATA
 232351 AAGGAGCTTC GCTCCTTGGG GGAAGGTAGC AAAGAGAAGT TCCGAGACAT
 232401 TCTCAGGTAG GGGGGCAATC GCAGAAAGGG TTAAAAAGGG GTTTCAGAC
 232451 AAAAATACCG AAACAGGAAG GTTTTGCTTT TTTTGCTCTG CTTCATACAG
 232501 ATGCATCCCT CCGCCCTTCT GGATTGAAA ATGGAGGCC ATGGTGTTTT
 232551 GATTGAACCG TTGCACGCGA TACATCCCAA GATTGGGTGT AGTAAGAGTC
 232601 GGCGATTCCG TATAGACAAG AGGAAGTGTG AGAAAGGCTC CACCATCTTC
 232651 AGGCCAGCTT GTGAGTAAGG GAAGGTGATC TAAGTTAACT GAGGACATAG
 232701 AAACAAAAGG AAAGCGACGG AATCGAGCTT TTTTGAGCCC TAAAGAGCTT
 232751 ATTCTTTTTA ATAGATCCCG AGATTTCAT AGAGAAGAAA GCTTTGGTGT
 232801 AGAAGAAATA AGGTGGGCAA CTCGAGCGAT GAGGTATCA GGAGCTTGAG
 232851 AAAAAAGTTG GTCTACACGA TGTTTTGTTC CAAAGAGATT GGTGAGGACT
 232901 GGAATGACG ATCCGATGAC ATTATGAAAA AGAAGGGCAG GGCCTTGATC
 232951 TTCAATAACA CGACGATGAA TCTCAGCTAA CTCGAGGTTA GGACTTACGG
 233001 GAGCAAAAAC ATCAATAAGT TGTTTTGTG AACGAAAAG AGAAATATGA
 233051 CGCCTTAAGA AAGACATAAA TATTTCCCTA TACTTAAGTT AAATTAAAAA
 233101 TTTTACTTT TAGCTCTTC GAGAACTTC TCTAATCCGA GCTTATCAAT
 233151 GTGACGAAGA GCGCTAGCAG AAATTTTAAG CTTAAGAAAA CGGTTTTCTT
 233201 CTGTAGACCA TAGACGCTTG GTCAACATAT TAGGGAAAAA TCTTCTTTTA
 233251 GTCTTCCCTG TCACTTTCAA ACCAATTCCT TTTTCTTTT TAGCAATACC
 233301 TCGAAGTGTA TAGCTATAAC CACGGCGAGG TCTCTTCCT GTAAGTGGGC
 233351 ACTTTCTTGA CATATTCTTC CTATGAATTC TCTAACATCC GCTCAATAAA
 233401 GCAACTCTAT GAAAAGGAAT CTATACTAGA GCTTTTCGAA TATATGGGGA
 233451 AGAGGTTTTT TTAGAAAATG TGATTGTCTG TTGTTGATAA TGAAGGAATT

233501 CTTTTAAGAA AGACGGTCGT TCCAGGAAAT TATTATAAGG TTCTTACTAT
 233551 AAAAAGACTT AAATGTTTTA TTGCTATCCT TACAGTCCTG TAAGGATCTT
 233601 CTCAATGTAA CCATTAAATT TTTTATGAAT AGCGAGTTCT TCTAAGGAAG
 233651 GCCGAAC TCG ATACGACCAA TTCTTTT TAG AAATTGTCCC AGGTGTATTA
 233701 ATGCGTTCTC TTTGTAGATT TTTTGATACT AAATCAGGGC AGAGGGCGAG
 233751 ATAATCGTTA AAGAGGTTGA TATGAAAGAT AGATGCTGAT TCATGAGAAA
 233801 GTTTTAAGAT GTCTATTTGA GTTCTGTAG TCAGGGTTTT TTGAAAAGGA
 233851 AGATGTAGAA ATTTAGCAAA TTGCTTAGCT TCCTTAGGTG AATTGAGCCA
 233901 CCATTGGGCA AACGTATCAG AGTCGTGGGT AGAGAGAGTG GTCACAGAAA
 233951 GTGGATTATA ATCTTTTAGG GGAATGAAGG CACTGTCGCT TTCCCAGTTG
 234001 CGTTCCCATC GTGGAATCCG GGTCCACAG ATTCCCTAAGT GTGTTAATGT
 234051 CGTTTTGACG TCTTGGGGTA TAATCCCTAA ATCTTCTCCG ATAGGTAACA
 234101 TAGAAGAGGC TCCGAGCATA GTAGAAAGGA TCTCCGTGCC CTGCTTTATA
 234151 TAGTCTTTAG GATTGTCTGG AATGAACCTT CCTCTTCTG AAGAATCCCA
 234201 AATCCACAAA CGGAAAAATC CTATAATATG ATCTAAGCGA TAGACGGAAT
 234251 AGAAGTTTTG AGCATATCGC AGACGCTCTT TCCACCAAAT GTAGTCGTCT
 234301 TTGGCAAGTT GTGAAAAATT ATAAATAGGC AGATGCCAGT TTTGTCCTTC
 234351 AGAATTGTAG AGGTCAGGAG GAGCTCCTAC AGACCTTGAT GAAGAAAAGT
 234401 AGTCTCGGAA ATACCAAACA TCACAGCTAT CCTTGCTAAT AAGAATAGGG
 234451 AGGTCTCCTT TAAGCAGGAC GTGGTGTGA TCTGCATAGG CTTTCACTTC
 234501 GCAGAGCTGT TGGTAACAGA GAAACTGTAG ATAGGAAAAA AAGAGGACTT
 234551 CATCATGGAA TTTTTTAGTT AAGTCCGAA AATTCTCCTG ATCTGTGAGC
 234601 GACTTCGGCC AGTTATTAAT AGGTCTCCG TGCATATGAT GTTTGATTGC
 234651 ACGAAAGGTC CCATAGGGAT AAAGCCAATA GCGCTCGCTT TCTAGAAACT
 234701 CAGAAAAATT TGAGTTTCCT TCGAGGGAAG ACTTGCAACA TTTTGGTAG
 234751 TACTCTCTTA AGAATGCCCA TTTTTTTTCT TTAACCTGAG TATAGCTGAC
 234801 TGATGGAGTC GAGCATAACT CATGCATATC TTGAAGTTTC TTGGCAACTT

234851 CAGGGATGGT ATCGATATTT GGAAGAGAGG ATAGGGAAAG GAATAGGGGA
 234901 TTCAGGGCTA CGGAAGAGAT GCTGTTATAG GGAAGCGTAT CTTACCAGT
 234951 ATCATTTAAA GGGAGAAGCT GAATAACGCT GAAGCCCTGT TTTTGGCACC
 235001 AAGAGATCAG AGGAATGAGA TCTAAAAATT CACCGATTCC ACAGCTATTT
 235051 TTTGTGTGTA TTGAAAATAG TGGGAGATAA ATCCCGTGTT TAGGAGAGGT
 235101 TCCTATAAGT TTCCAAGCAT GTGCTGAGGG TGAGTGTTTT GTGTATTTTA
 235151 AAACATTAC GCGACGTAGT AGATTCCCAA AACATGAAGG TCAGGAAGAT
 235201 TCCCTGTTCT TAGGGATTCC ATCCAAATCT TAGCATGCAA AGAAAAGAGC
 235251 TTGAGATACT CGAATAACTT TTCTCCATTT AGGTAGTTCA TACTCAGGAT
 235301 ATCTGAAAGA TAGAGCTGTT GAGTGACCTC ACCGTAGCCT AGAATTCCT
 235351 TGATGCTGGA TTGGAACGAG CCATTACAG AGAGAGCAGC TTTGAAAATA
 235401 CGCTCGCGAA ATACGTTTTT AGGAAGAGTA CCTAGTAGTG TCGATACTGC
 235451 AAGATCTCCG GAATTTCCAT CTTCTTCTAT TTGCACAGGG ACATGGGTAT
 235501 CGCTGAAACG GATTAAACAA GAGTTATTTT TATCTGGAGC AAGTGTCTGA
 235551 TTTAATAGGG GTGCTAAGGA TTCTAGTAAT TGCTCGTATT GGTTTTGCAT
 235601 GGCAATCCTT TTTTAATCAT GACCAAGGAT AGGGTTTAGG GAAGTCTGAT
 235651 GCTTTGGGAT AATCTTCATT GTTTATATTT ACAGCATCTA AAGCATTAGC
 235701 AATCATAGCT CCTAATTGCT GACGTTTGTC TGCTGAAGAG AAAAGGCGTG
 235751 ACGACGTTTG ACGTAAAGCA GAAAAGAATA AGTTCAAGAC ACCGGTCACA
 235801 GAATCAACAT CGTCTCCTAT GAGATTGCGG ACTTCTCGTT CTACTTTAGA
 235851 TGCTGTTGGG AACTTATCGT TAATGATTTT ATGGTAGGAC TCAGCTACCT
 235901 TCACAAAGTT TAGATCAGAA GGAGTTTGGG TTCCCTCAGC TTTTAAGCTA
 235951 TCGAGTAAAA TAGGAACGCG ACTTTCAAAG TAATCGTACG AGGTAAGAAC
 236001 TGCTTGAGG TTACGAGTTT CTGTCATGAG AACTTGAGT TGCGCACTGG
 236051 GTACGTAGGG ACCCTGCCTT TTTAATTCTG TTGCCATTCC TTTCATTAGA
 236101 AAGGAGCTGA CAATAGCCAT ATCTTGAGT GTATAGCGGT CTTGAAGCAT
 236151 AGAAAGTAGC TGATCACAGG TATGTGTGTC TCCAGTCACT TCTAAGTACA

236201 AAGAGCGAAG CCCTGAAGGA GAAACATTCA GTTGGTCTGC ATATTCTTGA
 236251 GAGGCAAATA AGATGTTTTT CGCACCAATA GCAGTTCGTC CGAATTGCTC
 236301 CGTATGAGTA TTCCTTGCTT GGATAAGCGC TTCTTTTAAT TTACCTTGGG
 236351 AGGGTGGAGT CGTTTGAACC AGGTAGTCCA AAGCTGTGGA TTGCAGAGCT
 236401 GGGTCTTTAA TTTTCTCTTG TACAAGAGCA AGAATGTCTT CTGGAGAAGC
 236451 ATCGTCTCCT ATTGCATCAC GCAGGCCGCG AAGTTCTTGA CCAGAGATTT
 236501 CAGAATTCCC AGAAGCATAC TTATCAGCAA GATCTGTGTC AGGCTTCTCT
 236551 TCTGTAGATT CAGATTTTTT CTCAGCCTTT CCAGCTTCTC CTTTTTTCCG
 236601 AGATTCTAGA GTTTGAAACT TCTCTTCCTT TTTTTTCGTG CGTGTGCTG
 236651 CTGCGGGATT TGTCAGGTCC TGAGATTGTT GAATCATGTT CATCTCAGAA
 236701 CCTTCTTGGC TGGCTACAAC TTCTGCTGCA TCTGCTTTTG CAGCTGCAGC
 236751 TTCTACAGCT GCAAGGTGA CACCCTGAGT GCCTCCTAAA CCACCTGTGC
 236801 CTCCTGATGC TGCCATATGC CTCCTATGAG CGACAACGTA TCAATTAGAA
 236851 AATCTGAATT CTTCTTAAAG GCTGGATGCG GATTTCTGGT AGGATTTCTT
 236901 GATAAGAAAT CACAGCAATG TCAGGGAATT CTGTTTCTAT TAATTTTCGT
 236951 ACATATCTTC TTACATCAAT TGCTGTCAAT AATACTGGTG GTTGGCCTCC
 237001 TGCAGGTGTT GCGGTGATCG TATTCCTCAT AGATTTTAAA ATTAGGTTCA
 237051 CAGAATCAGG ATCTAGAGCA AGGTAAGAAC CTGCCGATGT CTGTTTAATT
 237101 GCTCCACGAA TCATCTCTTC AATTTCTGGA TCTAAGAGAT AAACAGAAAT
 237151 TGCTGATTGT CCTTGAGAGA ACTTGAAGCT GATATAAAGC TTAAAGAAG
 237201 ACCGTACATA TTCTGTAAGC AAAACTGTAT CTTTCTCAGT TTGCGCCAC
 237251 TCGCTCAGAG ATTCTAAGAT TGTACGTAGG TCTTTAATTG AGATTTGCTC
 237301 TTGAACCAAT CTCTTAAAGA TTTCCGTAAG CTTTGTGCAAT GGAATAAGCC
 237351 TTGTGACTTC CTTCACTAAG TCCGGAATG AACGTTCCAT AAATTCGATC
 237401 ATAGAACGTA CCTCTTGAAT TCCCAAAAAC TCTTGAGAGC TTTTATGGAA
 237451 AAAGTACGAA AGATGGAGAA TGATCACTTC GAGCGGCGTC CAATATTTAA
 237501 TTGCTGCCTT CTCTAGAATA GCTTTTGCAT CTTCACTAAC CCAAGCTGAA

237551 GGAAGACCCG CAGCATTCTT ATAGGTAATG AAAGGTAGAT TATAACGGCT
 237601 GAGATTGTCC TCCACCTCAT TGGTTAACAC ATGGTGCGGA GGAATTTTTC
 237651 CTCGCACATA AGGGACTTCA TTAAGCAGAA TCATATAATC GTATCCTTCT
 237701 AAAGAAGGGG AATCTGTGCG AACATGAATG CCAGGGTATC GGATTCCGAT
 237751 ATCCTGATAG AGAGCTTGCC GCATTTTAGG AATCATATCA TCAACAAAGC
 237801 TTTGTCCTGA TTTTGTCTTG TGTGGATAA GCTTAGAGAG ATCTTTTCCA
 237851 AGTTCTAGAA TTACGGAAG AGTTAGAGAA TAGTCATCGG GATTATCCCC
 237901 AACAGTAGCA GCGCCATCAC CAGCAGCCCC TACGGTTGTT GAAGCTCCTG
 237951 AGCCACCACC TTTTTTTCCT GCCGCTGATT TCTTAGTCAG TAGGAGAATC
 238001 CCTAAGGCAA CGAAAATTAA TGCTAAAATG GAGAAGGACC ATAGAGGGAA
 238051 GCCCTTGAAG AAACCAACCC CTAAAGTTGC AGCACCTGCA AGGAGTAGTG
 238101 CTCGTGGTTC TTTAACGAGC TGAGTAGAAA TCTCTTTACC CAAGTTCGTA
 238151 TTTTGTGAC TCGATACACG AGTCGTGACA ATACCCGCTG TCAACGCAAT
 238201 CAAAAGAGAA GGAATTTGAG AGACTAAACC ATCTCCAATG GAGAGAAGAG
 238251 TGTAGACGTG AGCTGCTTGA GCGAGGTCCA TGCCGTGCAT AGCCACCCCA
 238301 ATCGTCAAAC CGCCAACAAT GTTAATCAAA GAGATAACGA TACCAGCGAT
 238351 AACGTCTCCT TTGATGAACT TCATGGCACC GTCCATGGCT CCGTAGAGTT
 238401 CACTTTCCTT TTGATTTGA GCCCTTTTAT CACGAGCTTG TGTGGCATCA
 238451 ATCATAACCAG CTCGTAAGTC CGCATCAATC GCCATCTGTT TACCTGGCAT
 238501 CGCATCCAAT CGGAATCGGG CAGCAACTTC GGCAACACGC TCGGCACCCT
 238551 TAGTTACTAC GATAAACTGA ATGATTGTAA TAATGAGGAA GATAATGAAC
 238601 CCGACCACAT AGTTCCCTCC AACCACGAAG TCTCCGAAGG CCTGAATGAC
 238651 ATGACCCGCA TACGCTTTAA GGAGAATCTG TCGAGAAGAG GAAATATTAA
 238701 TCCCCAAGCG GAACATCGTA GTGATGAGGA GCAACGAGGG AAAACAGAC
 238751 AGCTGCAAAG CACTTGGAAT ATAAAGAGCC ACCATCAATA AGAATACAGA
 238801 GATCGATAAG TTGATGGTGA TCATCAAGTC AACGATAGGC GGAGGCAAAG
 238851 GAATAATGAT CATTAAGACA ACGCCCATCA TCCAAAGAGC AAGGATTAAG

238901 TCGCTGGACT TATTGATCAT GTTTAAGGCG GTATCGCCAC CAAGTGTTCCT
 238951 GCTGACGAAA TTGAGTAGCT TATTCATTAT AAATGATCAG GTTGGTTAGT
 239001 ATTTTTATTA TTAGGATTTT GCGCATTCAG TGAAGTGATA TAGAGTAGAA
 239051 TTTCTCCAAT AGCTTCGTAA GTAGATTCTG GAATAAATTT TAATTCCTTC
 239101 CCTTCATCCA AAAGCTGATG TGCTAAAGGT ACGTTTCGCA TAATGGGAAT
 239151 TCCGTACTTT TCAGCTTCAT CAAGTATCCT TTTAGCTCGT AAGTTGATGC
 239201 CCATGGCAAT GATCCAAGGT GCTTTATATT TTTCAGGCAT GTAGCCAATA
 239251 GCAACAGCAA TATCTTTGGG ATTAGAGACT ACGGTGCTTG CATGTTTCAC
 239301 CTGTGATGAC GAGTCTTCAT AGGCAATTTT TTGAGCAATT TGTCGACGAC
 239351 GGCTTTAAT CTCAGGATTT CCTTCCGTGT CTTTAAACTC CTGCTTAACC
 239401 TCAAACCTCT CCATCTTTAA TTCTTTAGCG AAATTGTGGC GCTGATAGAC
 239451 AAGGTCAAGA ATCGCAACAA TCAAAAAGAA AATTCCTATC GAGGTTACTG
 239501 CTTTATAAAA AATTTCTTTG AAGATTGAG CAGTAATTAT AGGAGAGACT
 239551 CCTGCAGTTT CTATAATTAA AGAGACTTTG CTTTTTAACG TTATGTATAA
 239601 AATTAAGGCT GCTCCAAAAA TTTTAAAAAT CGATTTGATC AGCTCTATGA
 239651 GAGTCTTTAT TTTAACTTTT TGTTTGATGT TCTCAATAGG GTTGAAC TTC
 239701 TTGATATCTG GTTTAAAAAC TTCGCTAGAA AATGTAGGAC CAACGATAAG
 239751 AAAACCTACA ATGACGCCAA CAACAGCAAC AGCTCCCAGT AAGGGAAGTG
 239801 ATGCTGTTAA AATAAGCATA AGACAGTTCT TAAGATAAAA TAAGGTAATT
 239851 ACAGGATCAT GGCGAGTGGG AGCTTGTGAG AGCATGGAAA CCAGAAAGCC
 239901 ACCTAAATGC TTGAAAAAAA AGGTCGATAG GGAGAAAGCC GTAAACATAG
 239951 AGACGATAAA GGTAACCGCA GAAGGAAAAT CCTGAGATTT TGCTACTTGA
 240001 CCTTTTTTCC GAGCATCTCT AAGTCGCTTC GGCGTGGCCT TTTCTGTTTT
 240051 TTCACCATG CTGTTGCCAA GGTTCGATTA AGAGGACTAT ATTCGGAATT
 240101 ATCTGCAAAT TACCTTATGC TCTTGTTTTT CTCAAGAGTG AAGCCCAAAC
 240151 GTTTTGTTCT TAGTAAGATT ATAAATGAAT CTTAGAGATA AAGAAGAATA
 240201 GTAGAAAAAA AGAACAATAA CGAAAAGTTG TTGAATTTAG CAATACTTTA

240251 GGTCAC TTCA AGATGG TTCT ATATGA AGAA AGTATT AGGT AGGGAAAA CT
 240301 TCAATCG ATT ATAGTG GTTT TTTATC ATTT TCAAAG CACT CTTGAT TTTG
 240351 AAGGTACT AT TTTTCT AGCT ATTGTG CTTT TTTGAA ACAG AAGAGT TGTC
 240401 TTTCTA ATAT TCGAAAA AGC ATGTCATT AT TTTTAT TTTT AGATGT CTTA
 240451 TGAGTC ATAC TGAATGT GGA ATTGTAG GGC TTCCTA ATGT AGGAAAGT CT
 240501 GGCTTATT CA ATGCTTT AAC AGGAGCT CAA GTTGCC TCCT GTAAC TATCC
 240551 GTTTTGT ACT ATCGAT CCTA ATGTGG GTAT TGTTC CTGT ATCGAT GAAA
 240601 GACTGGA AGC CTTAGCT AAA ATTAGCA ATA GTCAGA AGAT CATCTAT GCG
 240651 GATATGA AAT TTGTAG ATAT TGCAGG TTTA GTTAAG GGAG CTTCCG ATGG
 240701 CGCGGG TCTG GGAAAT CGGT TTCTCT CTCA TATTCG AGAA ACTCAT GCTA
 240751 TTGCTCAT GT AGTGC GTTG TTTGAT GATC CAGACG TTAC ACACGT TTTCA
 240801 GGAAAAGT CA ACCCTGT TGA GGATATT GAA GTTATCA ACT TAGAGCT CAT
 240851 TTTTTCTG AC TTCTCCT CAG CAAAAA ATAT CCATAG CAAA TTAGAAA AGC
 240901 TAGCCAA AGG AAAGCG TGAA GTAGGAG CTC TCTTGC CTCT ATTTGATA CA
 240951 ATTATTG CTC ACTTAG AAAA GGGGCT GCCG CTACGT ACTT TAGAATT AAC
 241001 TCCAGA ACA AATTGT GGCAT TAAAGC CTA TCCGTTTT TG ACCATGA AGC
 241051 CTATGTTTT A TATAGCT AAT GTTGAC GAGA GTTCTCT ACC AGATAT GGAT
 241101 AATGATT ATG TTGCCG CTGT TCGGGA AGTT GCTGCA AAAAG AAAATT CTAA
 241151 AGTGGTT CCT ATCTGT GTTC GTATAGA AGA AGAAAT CGTT TCCTTAC CTA
 241201 TTGAAGAG CG CTTAGA ATTT CTTATG AGCT TAGGTCT TGA AAAATC AGGA
 241251 CTTCATAG AT TAGTGCGT GC TCGGTAT GAC ACTTTAG GAC TGATTT CTTA
 241301 TTTTACT ACA GGTCCT CAAG AATCTCGT GC ATGGAC AGTG GTTCGAG GGT
 241351 CTTCTG CTG GGAAGCT GCT GGAGAA ATCC ATACGG ATAT TCAAAAG GGC
 241401 TTTATTC GTG CTGAAG TGAT TACTTT TGAA GATATG ATAG AGTGTC AAGG
 241451 TCGTGC AGCT GCTCGA GAAT TAGGGAA ATT ACATATA GAA GGACGT GATT
 241501 ATATCGT CCA GGACGG TGAT ACTATG CTGT TCCTTC ATAA TTAAAG GAAC
 241551 CCTTTG CAAA CCAATCT TGA GCATCC AAAA TGTCTTTT TC AATTGCT CGT

241601 ATTAGAGTTT CTTTTGATTG AACTTTTTTT TCTTCTCTAA GAAATTTTCT
 241651 CGGGATAATG CTCACCTCTT TGCCGTATAG ATTTTCCGCA AAGGAAAAGA
 241701 TATGCGCCTC TGCATATAAA GACTCTCTTC CAAAAGTAGG GGCAGTTCCT
 241751 AAATTCATAA CACCCTGACA GGTAGTGCTA TCATAACGTA TTTCACAAGC
 241801 ATAAACTCCT AGGGGAATTA AACTTTCTTC TCTAGGAAGA TTTATAGTGG
 241851 CGAATCCTAG AGAACCTCCT ATTCCGAGC CCTCGGTTAT TTTCCAGAA
 241901 ATGGCATAGG GATGACCCAA AAAACGATGA GCACATTCAA GATTCCCTGC
 241951 GGACAGAAAC TGGCGGATTG CTTTGCTGGA GACAACTATG TTATCCATAC
 242001 GGTAAGGAGG AATCTTGATG ACCTCTATAC CTAACGGCTT GCCTATAGTA
 242051 TCGAGAGCCT CGGTATTGCT TTGCTGTTCT TTCCCTATGC AAGAATCATA
 242101 ACCTAAGATG AGGCGTTTGC ATTTCAAGTT ACGATGTAAC AAAGTAAGAA
 242151 ATTCTTCTGC CGATTGATC GCAAAATTTA AATCAAAAGT AAGGACACCT
 242201 AACCAGTCTA TGGGAAACGT TTGCAATAAT TGGAGGCGCT CTTCTTTTGT
 242251 ATTGATGAGT TTCGTGTGAT TTAAAGAAAG TACCGTTTGA GGATGAGAAT
 242301 CAAAGGTAAT AACTCCACTG GATCCAGAAT AGGAAGTAAG AATAGATAAA
 242351 AGATTGCTAT GCCCTAGATG ACATCCGTCG AAAAAACCTA CAGTTACAGA
 242401 ATCTACAGAA AACGAAGACG TTAAACTATA GGCTATTTCC ATGGGCATCT
 242451 CGTAGGTAGG GAGAAATATC GAAATCGGGG TGGTCTAATA GATTCCCATC
 242501 AATACATTCA TCTATAGAAA AACGGCCACT GCGTAAACGG CGTAGCTGCT
 242551 CAAGATAAGC TCCACAGCCT AACATCGTGC CAAGCTCATG AGCAATGCTG
 242601 CGAATATAAG TTCCTTTGCT ACAAGAGACT ACAAATGCA ATAAAGGGTA
 242651 CTCATATTTT GTAATCTGCA AGTGAACCTG AACTGTAGAA TGGTGACGTT
 242701 CTATAGATAA ACCTTTTCTA GCATATTCAT ACAGCTTTTT CCCTTGGACT
 242751 TTTTTAGCGG AAAACATGGG AGGCAGTTGC TGGATCTCTC CTTGGAAATA
 242801 CTCGGCAGCT GATAATACTT CTTCGAGACT AGGAATCTTC TTAGATCTTC
 242851 CTACAACTTT GCCGTCGCAA TCATAAGAAT CGGTAGTTGT CCCTAAATGG
 242901 GCAATTGCTT CGTATTCCTT GTCTTCAAAA AGTAAAATAT CAGAAAGTCT

242951 AGTAAATTTA CGGCCAATCA ACATGACCAT AACGCCAGTA GCGAAGGGAT
 243001 CTAAAGTTCC TGCATGACCA ATCTTTTTAA CGCCTATTAA CTTGGTTAGA
 243051 GCGCGGATAA GGCTAAACGA AGTTCTCCCT TGAGGCTTGT CTACAAGAAG
 243101 AATGCCCTCT TTTAATTCTA CTGCAAGATC CATAGTCATG TCTTTAATAG
 243151 TATTCAAGTT CCCAAAAAT ATAGTTTATT AACTCTTTTC TTTCTCTTGA
 243201 ATCTGCCAAA GCAGGTTTTTC TATATAATCT TGAGGTGAGA AAATATCATC
 243251 GAGATAAAAA TGAAGTTCTG GGAAATATTT AAGGACGACA TTTTTCGAAG
 243301 CTCTATGAGC GATAAAACCA GCAGAGACTT TTAAAGCTTC TAAAGCCTCT
 243351 TCCTTGGTAT TCTCATGAGG CATTACAGAT ACATAAACAC GTGCAGAGTG
 243401 CAAATCCTTA GATAGAGAAA CACGAGTTAC CGTGATCCAA AGATTAGAAA
 243451 TCTTGGGATG CTTAACATCT TTTAAAATTA CCTTTGCAAT GGCTTCTTGT
 243501 AATAAAGCAT TTACCCGTTT AATACGTCTA TTTTCTGTCA TACAGTACTT
 243551 CAAGTTATAG TTTTGTGGA TGATAGATAA CTTCATAACA TTGTAGGACA
 243601 TCACCTATTT GAGCTTGCTG GTATCCTTCT AACAAAATTC CACACTCTAA
 243651 ACCTTTGCGA ACTTCTTTGA CATCTTCTTT AACACGTTTT AATGAAGATA
 243701 ACGTACCTTT CCAAAGGATC TCTTTATTAC GTAATACTCG GACTTTATGA
 243751 TTGCGAGTCA TAATTCCTTC AGTAACTATG CAACCGTAAA TAGATCCTAC
 243801 TTGTGAAGAC CTAAAGATTT CTTAATCTC AGCAGAACCT TCATCTTTTT
 243851 CTTCAGCAAT AGGATCTAAT AGAGAAGTCA TAATTTCTTT AATTGCATCA
 243901 ATAGCATGAT AGATGACGGT AAATAGTTCA ACTCGGACTC CTAAGCTCTT
 243951 AATTAAAGGT TCCGCATGAC TTTCTATTCC TGTATGGAAA CCGATGAGAA
 244001 CTGCTTTAGA GGCGGCAGCT AAACGAATGT CTGATTCTGA AATTTCTCCT
 244051 ACACTGTTTG TTAAAATTTC AACATCTACT TTTTCTGATT TAATCTTAGA
 244101 TATTGAACTG ACCAAAGCTT CTATGGAACC TTGAACATCA GCTTTAATCA
 244151 TAAGCTTAAG AGTCTTTTTA TTCTGTAACA TAGAATCAAA GTTAGGCCGC
 244201 TTCTTTTGCT GTAAAGCAAA ACGCTGTTGT CCTGCGGATC TAGCTTCAAT
 244251 AATGTCTCTA GCCGTTTTCT CGTTTTTCAC GACGAAGAAA GGATCGCCAG

244301 CTTTAGGAAT GTCCGATAGA CCTGTGATCA ACACAGGAAT AGATGGCCCA
 244351 GCTTCTTTCA TCAATTCATT ATGTTTCGTTA TGCATAGTTT TCACTTTGCC
 244401 ATAACAATCA TTGAAGACGA GAGCTTCGCC CAGTTTTAAG CTTCCATTTT
 244451 GAATCAAAAC AGTCGCAACA GGTCCGAGAC CCTTGTGCAG TTCTGATTCA
 244501 ATAACAAGTC CTCGAGCACG TGCTGAAGGA TCGGCTTTTA GCTCCAAGAC
 244551 TTCAGCTTGT AAAGCTAACA TCTCTAAAAG TTCTGAAAGA CCTTCTCCTG
 244601 TTTTTCGCGA GGTATTTACT GTAACAGTCG AGCCTCCCCA AGCTTCTGGC
 244651 AATAGATTGA TTTCAGAAAG TTGTCTATAG ATGGTTTCGG AATTAAAATT
 244701 AGGCTTATCA CACTTGTTGA TAGCTACAAC AATAGCGATA TCAGCAGCTT
 244751 TTGCATGTTT AATAGCCTCT AAAGTTTGTT CTTTAATTCC TTCGTCTCCA
 244801 GCGACTACAA GCACAACAAT ATCACAAACT TCAGCTCCAC GGGCTCGCAT
 244851 TGCAGAGAAA GCTTCGTGAC CAGGAGTATC TAAAATTGTT ATGTCTCCCA
 244901 CTGGGGTGGA GCAGCAGAAG GCTCCCATGT GTTGGGTAAT CGCTCCAGCT
 244951 TCTGTTGCAG CGACATTACT TTTCTTAAG GAGTCAATGA GTGTTGTTTT
 245001 TCCGTGGTCG ACGTGACCCA TAAACGCAAC AATAGGGGAG CGAATCACAA
 245051 GCTTGCTGGG ATCTGTAGAT TGAATTTCTG CTCTTACAGT GTCATTGCTT
 245101 AGGCACAAC TATCTTGCTC AGAATAGTCG ATGTCAATTG TACATCCAAA
 245151 CTCTAAGCCA ATAAATTGTA CTGCAGTTTC GCTGTCTAGA ATATCATTGA
 245201 CTACATAGGT CATTCCATGA ATGAATAACT TTTGAATGAC TTCTGAAGCC
 245251 TTGAGCTTCA TTTCTGCTGC CAGATCTTTG ACGGTAATTG GCAAGGAAAT
 245301 TTTGATATGC GTAGGTCGCT GGATAGAGGC TTCGTCATAG TGTTTTTTAG
 245351 GCTTATAAAC ACGTTTTTTT CGCCATCTGT CTTCTTCTCC GCCTTCATTT
 245401 AATCCGTAAC GATCTCTTCC TGTAAGAGCC TTTAGGCTTT CATCAGATTT
 245451 CTTAGAACGA TCACGAAAGT CGGTAAGATT TTTCTTCCCA GCATCTCTTT
 245501 TTGGACCAGA AGCAGGACTA CGGTTAGCAG GATTGAATTG TTTTCTCGA
 245551 TTATTCTGTT CACCACCTTC AGATGTTCCA GGTTCCTCTG TTTTATCTGA
 245601 AGCAACGGGC TTTGTGCTTT TCGAGCCAGC TACGACTTTC TCTTCCTTGG

245651 CAGGAGCCTT GAATGTTTTT GCTAGGAGAT GATTGATATG CTTTCCTGTA
 245701 GGGCCGAACCT TAGATTTAAT CATTACGACG CTTT TAGGTT CAGCAGGCTT
 245751 CACAGGCTTA GGCTCTAATT CTTTTTCTTG AGGTGGGGTT TCGGGCAATA
 245801 CGGGTTGCTC AGGAAGAACT TCAGCAACTG GATGAACCTC AGGACTTTTCG
 245851 TCACAAACCT CATCGACTAC TTCTAACTCA GGCTCAGGAT CTGCTATGGA
 245901 GACTGGAGCA GGTT CAGATG TATCCACTGG AATATGAGCA GAAGACTCTT
 245951 CTTCCGATGA TGAGAACGAC GAACGATTTT TAGCACGAAT GCGACGTGAA
 246001 GTAGACTCTG GTGAAGCTTG TTCCGCACTT GCCGTAGGGG TAGAAGTTGC
 246051 GGCAAGAGCT ACTTTTACAG ACTTTTCTTT CGCAGAAGGT TTTTCTGAAG
 246101 AAGATTTAGC TTCAGAAGAT CCTGCTTGGG CAAGTTTTTG TTTTAGCTTA
 246151 TCCAGCCCCG CGGCTTTCGT TAATTGAGCA TTCTTAATCT TCAATTT CAG
 246201 GTTTTTCGTC AACTTTACTT TCTCCATATT TGCTGACTTG CTCAAGGATC
 246251 TTATAAGCAA GCTCTAAACT GATCCCAGGA ACAGATGCCA GATCATTAGC
 246301 ACTCGCTAAT AATACTCTTC TAATTGTGTC ATATCCAGCA TGTCTAAAT
 246351 TTTGGATGAC TAGCTTACTA ATCCCTTCCA TTTCTAAGGG TTGATCTAAG
 246401 TGCGGACTAT CGAATTCTGC TAATTGAAGG CGTTGAATTT CTAGCAACTT
 246451 ATTGTACTCA CTCATACGTT GTA CTTCGAG CTCGTAGTCT AGAATGTGGC
 246501 TAATTAAACG AGCGTTAATT CCTCGTTTAC CAATAACAGT AGCGTAGTCT
 246551 GCATCATTA A CGACAATTGC AATCACTTTG TCGTCTTCTA AAATAGCAAT
 246601 CTTTTGGATT TCTATTGGAT AAAGAAGATT CTGTAATAAC TCTGTAGAGA
 246651 CGGGGGAGTA ATTGACAATG TCAATTTTCT CATCGTTCAA TTCTCGAATG
 246701 ATATTTTTTA CTCGAGAACC TCGCATACCT ACAAAGCTC CAACAGGATC
 246751 AGTTTTAGGG TCTGACGATC TTACAGCTAG TTTCGTGCGG TACCCAGCTT
 246801 CACGAGCTAT CTTAACAATC TCCACAGAAC CTTCTTCTAG TTCTGGGACT
 246851 TCTTGAATAA ATAATTGTTT AACAAATTCT GCGTGACTAC GACTGAGGAT
 246901 AACTTCCGCT CCACCATTTT CAGACTCTTG AACTTCATAG AGTAGGGCGT
 246951 AAATTTTATC ACCGATCTTA TGTTTTTCTG TTTTAGGATA AAACCGGTA

247001 GGAAGAATTG CTTCAACTTT TCCTAAGTCA ATAATTAAAT TAGAACCTTT
 247051 AGCAAAACGT TTGACAACAC CAGATAAAGT TTCATTTACG CGATGGCGAT
 247101 ATTCTTCATA AATAACGTCT CTCTCAGCAT GTCTTAGCTT TTGACCGATA
 247151 ATTTGTCGTG CTGCGTGAGC AGCTATCTTC CAAAATTATC AGAAACAAAA
 247201 GGGACATCCA TGTACTGACC AATCTGACAG TCCGGATCGT ATTCTCTGGC
 247251 TTTATCTAAA GGAATTTCTT TGCTAGGATT CTGACAAATT TCTACTATTT
 247301 CCTTTTCACA AAAGACTTCG ATGTCACCAG TACGAGAATT AATGTTTACA
 247351 GATATGTTCTG CGTCATCTCT TAAGGTTTTT TTAGCAGCAA TTTTAAAGC
 247401 AGATTCGATA GCTCCTATAA TAGTAGAGCG CTGAATCCCT TTTCTTTCT
 247451 CCATGTAGTC AAAAATAGCT ACAAGATTTT TATTCATTAT ACTCCTCTTG
 247501 TAAATAAGGA AATCAGTAAA TAGCTAAATT AAGGACGTAT TATTTAGAAC
 247551 TAAATAATAC GATCCTCTGC ATTACCAGCA TTAGATGCTA TTTTCCTTTT
 247601 TTCTTTCTCT CTTTAGGGCC TTGAGAATCC TTGAAATCTA ATTCAGTCCT
 247651 AGAGTCTTGA TCATAAGCAT TGTCAGCTAA GTATTCTTTT ACAGAAAGAG
 247701 AAACTTTTTT ATGATCTGGA TCTAGCTTAA TTACTTTTGC AGAAACATTT
 247751 TCTCCAATGG AGATAATATC TTCAATTTTT GCAAAGGGCT TGTCAGAAAG
 247801 TTCTGAAACG TGAATCAATC CTTCAATCCC GTTTTGTAGC TCAACAAAGG
 247851 CTCCAAATGC AGTGATTTTA GTCACAACCTC CTGAAATTAC TGTGCCAGCA
 247901 GGGAACATAG CTTCAATTTT ATTCCAAGGA TTAGAACTTA ATTGCTTAAC
 247951 TCCTAAAGTA ATTTTTTTTAC TTTCTTTGTC TACTGATAAA ATAACAGCCT
 248001 CTACAGAATT TCCTTTTTTG AATAGTTCTG AAGGGTGAGA GACTTTTTTA
 248051 ATCCAACCTCA TGTCAGAAAT ATGAATCAGA CCCTCAATTC CTGGTTCTAA
 248101 TTCAACGAAA GCACCGTAAT TGGTTAAGTT CTTGATTTC A GCATGACAT
 248151 GGAGACCTAT AGGATATTTT TCTTCGATAT TGTCCCAAGG ATTACGTTCT
 248201 GTTTGCTTTA ATCCTAGAGA AATTTTTCCT TCGTCCTTCT GAATAGATAG
 248251 AACAAATGGCT TCAACTTCAT CGCCTTTATT TACGACTTCA CTAGGATCTA
 248301 CAATATTTTT CACCCAAGAC ATTCAGAAA TGTGAATTAG ACCTTCAATG

248351 CCCTCTTCAA TTTCAATGAA AGCTCCGTAG GGGAGAAGCT TCACAATTTT
 248401 ACCAAGAAGT CGTTTTCCAG GAGGGTATTT CTTCTCAATA TCTTCCCAAG
 248451 GATTATGCTC TTTTTGTTTG AGACCTAGAG CAACTCGTCC TTTTCTTTTA
 248501 TCTACGCTTA AAATAATTAC TTCCAACCTCT TGATTCAATT CGACCATTTC
 248551 GGAAGGATGT CGTATGCGCT TCCAGGTCAT ATCGGTAATG TGGAGAAGAC
 248601 CGTCAATACC ATCGAGATCT AAGAATACAC CAAAGTCAGT AATGTTTTTA
 248651 ACAACTCCTT TCGCGTATTC TCCGATAGAA ATTTGTTCAA TAAGTTCGGC
 248701 TTTCTTAGAG ATTCTCTCAG CTTCTAAGAG TTCTCTTCTT GAGACAACAA
 248751 TATTGCGACG TTCAACGTTA ATTTTAAAAA TTTTGAATTC ACAAACTTTT
 248801 CCGACATAAT CATCTAAATT TTTGATTTTC TTGTTGTCAA TTTGTGATCC
 248851 AGGTAGGAAG GCTTCCATTC CAATATCTAC AATAAGGCCG CCTTTGACTT
 248901 TACGTGTAAT TTGACCTTTA ACAATAGAAC CTTCTTCACA ATGAGCTAAG
 248951 ATGTATTCCC ATTGACGTTG TCGTGTGGCT TTTTCTCTAG AAAGGACAAC
 249001 TTTGCCCTCT TCGTCTTCGG CTTGGTCGAG ATAGACTTCT ACTTCAGCTC
 249051 CAAGCACTAA ACCTTCTGAA GAGTCTATGA ACTCTGACAT AGGGATCACT
 249101 CCCTCAGACT TCAGACCAAC ATCAACTACG ACAAAGTCTT TATTAATATC
 249151 AACTACGGTA CCTTTTAGGA TGGCGCCAGG CTGTATTTTCG TTATCAGATT
 249201 CTTCTTCGCT CGAAGTAATT CTGTGTGCCG TATAAAGCAA ATCTTTAAAT
 249251 TCGGCAACGT CTTCTGTGAG GCATTCTATA TTGTCCAGAA TTTTTTTAGA
 249301 TCCCCAAGTA TATTCAGCTT GTTTTGGCAT TTATTGTTAT TCTCCTAAAA
 249351 ACTACAAAGT CAAAAAAGAC AGTGTAATA TAAACCTTAA AAAAGGCAAG
 249401 ACCTCCCTTG CAGGAGACGA TTAGTTCTAT ATAATCGAAT TAATCGTAAA
 249451 ATCCTATACT TCCGTTCTAG AGATGTATAA GGCCTCTTTT TTAACCTTTTC
 249501 TATATTACCA GATTTTATAA AATACGTGTG GTAATTATCA TAAATGTGCC
 249551 TTGTCATAAG TAGAGAAACA GGATTTGGCG CATAGAAAAC ATATCGTTTT
 249601 TTTCATAAGC ACCCCATAAA TTCTCGTCCC AGATGGGCTA CATCAGCTTT
 249651 GTTCTTAAGC CTGATAGATC TTTTGGTTG CTGCTCAAGA ACCAAGTGCA

249701 TTCAGAAGTA GAACCATACT TATTGAGGGT TAAATAATAG GATGGGTCTT
 249751 GGGTTCTTAG CTCTAAGAAA ATTCTAAAGA TTAATCACAG AAATGTAGGT
 249801 GTTATTACTG GAAGATGAAG ATCTTGAGAT TTGAAACTTT AGTCTTAAAT
 249851 CCGTATAAAT TAAAACTAGA CAAAAGCTAT TCAAAAGCAA GTTTTTTGCA
 249901 TAGAGAACT GACGCATAGA CGCTAAGTGT TGTATTATTAG ATAGGCTAGA
 249951 ATTAAGTGTC TTTTTTCTAG TAGATCAAGA ACAGTGTCTT CAATTTCTAA
 250001 AGTTAGATAA GGAAATTATA AATGATTCAT TCCCGGTAA TTATTATTGG
 250051 TTCAGGTCCA TCTGGATATA CAGCGGCAAT TTATGCATCA AGAGCGCTTT
 250101 TGCATCCTCT TTTATTTGAG GGGTTTTTCT CTGGGATCTC TGGTGGCCAG
 250151 CTTATGACTA CAACAGAAGT TGAGAATTTT CCAGGGTTTC CTGAAGGGAT
 250201 TCTTGGGCCA AAACCTATGA ATAATATGAA GGAGCAGGCT GTGCGGTTTG
 250251 GGACCAAGAC ACTAGCTCAA GATATTATTT CCGTAGATTT TTCTGTTCGC
 250301 CCTTTTATTT TGAAATCAAA AGAAGAAACC TATTCTTGTG ATGCCTGTAT
 250351 CATAGCTACA GGAGCTTCTG CTAAACGTTT AGAAATTCCT GGAGCAGGAA
 250401 ACGATGAATT TTGGCAAAAA GGAGTGAAGT CTTGTGCCGT TTGCGATGGG
 250451 GCTTCTCCTA TTTTTAAAAA TAAAGATCTT TATGTGATTG GGGGAGGGGA
 250501 TTCTGCTTTA GAAGAAGCTC TTTACCTGAC TCGTTATGGA AGCCACGTAT
 250551 ATGTAGTTCA TCGTAGAGAT AAACGCGGG CTTCTAAAGC TATGGAAGCT
 250601 CGGGCGCAAA ACAATGAAAA AATTACATTT TTATGGAATA GCGAGATTGT
 250651 AAAAATTTCT GGAGATAGCA TTGTCCGTTT CGTAGATATT AAGAATGTTC
 250701 AGACTCAAGA AATTACAAGT AGAGAAGCTG CGGGGGTGTT CTTTGCTATA
 250751 GGCCATAAGC CAAATACGGA TTTTCTCGGA GGACAGCTGA CGTTAGATGA
 250801 GTCGGGCTAT ATTGTGACTG AGAAAGGAAC GTCCAAGACT TCTGTCCCTG
 250851 GAGTATTTGC TGCTGGAGAT GTTCAGGATA AGTACTATCG TCAGGCGGTT
 250901 ACTTCTGCAG GAAGTGGTTG TATAGCAGCA CTAGATGCTG AAAGATTCTT
 250951 AGGCTAATGC GATTGCAGTT GCTGTGGCAT ACTCTTTGCA GTGGCTTATA
 251001 GAGAGAATGA CTTTAGAAAT TCCAATTTTT GCATAGACAT GCGAAGGGAG

251051 GAGAACTTCG GGTCCGTGAG ATACTTTAAA GACTTCGATG TCTTTCCAGG
 251101 CAACAACGCT CCCTATGCCA GTTCCTAAAG CTTTGTGCTAC AGCTTCTTTT
 251151 CCAGCAAAGC GACCTGCAAA TGAAGGGATG GGATCGGTCT TTTCTAAGCA
 251201 ATATTTCTGT TCTGCTTCTG TAAAGATTCT ATTGAGTAGT CGATTGCCGT
 251251 GAGTTGCAAT TGCCTCGCGA ATGCGGCTAA TTTCAATAAT ATCGGTTCCT
 251301 ATATGAATGA TTTCCATAGA GTTCGCACTA ATTTCTTCA GGATTTTCCA
 251351 TTAGGGCATT CTGTAATTTT GAGAAGTAAG CGATGGCAGT TTTAATGAGA
 251401 GCATAGCCCA CTGTAGCTCG AATCACGAAT AAGGCTTGAG GATTGCTGAT
 251451 ATAACGACGA AGGACTAAAG AGAAGGCAAT CATAATGACC ATGATGAGAA
 251501 ACCTTCTAGA GATAAAGGTC TGCTTTATAT AGGCAATCCA TGTAATTTTT
 251551 TGTGAAAAGA ATTCAGAAGA TAGACTCAGC TGTTTCCTTA TAGTTTTGCT
 251601 TAGCAAGTAG CGATGTTTTA GGGATGCTAA TCCCCAAGAA ACCATAGATA
 251651 GAAGACAGTA GGTAAAAAAA GAAGAAAGGG GGTCTAGGGC TATTGCGGTA
 251701 GCTTTTAGCA AGAGTTTCAT TCCTGCTGAT ATCCACAAAA TACCAGGAAA
 251751 TAGTATCAGG AAATATTTGA TGTTTCTTGC CATAGTGCAT TACTTTAAT
 251801 TTCTAAATAC TCTATTTTAC TTGGAGTATG AATTTTTTGC TAACTGGTAT
 251851 TAAGATATGC GAGCTGAGAT GGCTGTGATC TATTGGGATC GCTCAAAAAT
 251901 TGTCTGGTCT TTCGAGCCAT GGTCTCTAAG ACTTACTTGG TATGGCGTCT
 251951 TTTTACTGT AGGGATTTTT CTAGCATGTC TCTCAGCAAG GTATTTGGCT
 252001 CTTTCCTATT ATGGTTTGAA AGATCATTTA AGTTTTTCCA AAAGCCAGCT
 252051 ACGCGTGGCT TTAGAAAAC TTTTATATA CTCTATTTTA TTTATTGTCC
 252101 CTGGAGCTAG ACTTGCCTAT GTGATTTTTT ATGGATGGAG TTTTACTTA
 252151 CAACATCCTG AAGAGATCAT TCAAATATGG CACGGAGGCT TGTCGAGTCA
 252201 TGGAGGCGTT CTTGGCTTTC TTTTGTGGGC GGCCATTTTT TCTTGGATAT
 252251 ATAAAAAAA GATTTCAAAA TTGACTTTTC TCTTCCTTAC AGACTTGTGT
 252301 GGATCAGTTT TCGGAATTGC AGCGTTTTTT ATTCGTTTGG GTAATTTTTG
 252351 GAATCAAGAA ATTGTAGGAA CACCGACTTC TTTGCCTTGG GGGGTGGTTT

252401 TTTCTGATCC TATGCAAGGT GTCCAGGGAG TTCCTGTGCA TCCTGTGCAG
 252451 CTTTATGAGG GAATCAGTTA CTTGGTCGTC TCTGGAATTT TATATTTTCT
 252501 TTCCTATAAG CGTTATTTGC ATTTAGGTAA GGGATATGTG ACTTCTATAG
 252551 CCTGTATTTT TGTGCTTTT ATTCGTTTTT TTGCGGAGTA TGTA AAAAAGC
 252601 CATCAAGGGA AAGTTTTAGC AGAGGATTGT CTA CT TACAA TTGGTCAAAT
 252651 TTTATCTATA CCTTTATTTT TATTTGGTGT GGCCTTACTT ATCATTGCT
 252701 CATTGAAAGC TCGAAGGCAC CGTTCACACA TATAGCAGCA TCTTTTTGGC
 252751 ATGAGACTTT AGAATTTTAC AGTCTATCAC AATAAGAGAC CTGTATTGAG
 252801 AGATTCTCCT TTGGGAGATA ATGTGGGTTT ATGATATTTT GACCTCTTAA
 252851 GTTTAATTTT TAGAGTTTAT CAAATGAATA AACGCACTTT GCTTTTTGTT
 252901 TCTTTAATTG GGATTGCTTT TG TAGGATGT CAAATATTTT TTGGTTATAA
 252951 TGAATTTTCG TCCTGCAAAA ATCTAGCAGA GAAACAAAGA AAGATTTT CAG
 253001 AACAGACGCT AGCTGCAGTA GAATCTGTAG GGT TAAGTGT AGCTTCATGG
 253051 GACACCGATG TAAACGGAGA AGAACATAAG AATAACTACG CAGTTCGTGT
 253101 TGGAGACAAG TTATTTTTAT TACATAATGG AGAAGCTGCT CAGTCTGTTT
 253151 ATTCTTCTGG GGAATCTTGG AGCTTTGTAG ATCACAAGTG TG GTTTTCGAT
 253201 AATATTCACT TGGCTTTATA TCGTCAGCAG GGTTCCTTCTT TCAATCCTAC
 253251 GAATACGGGA AAAGTTTTTC TTCCTACGAA TCATGAAGGT TTACCTGTAC
 253301 TAGTTGTTGA GTTTCGTAAC AATAAAGAGC CTTTAGTATT TCTAGGTGAG
 253351 TACGCACAAG GAAGAATTTT CAATAAAGAT AGCACGATCT TTGGTACAGC
 253401 GCTTGTCTTT TGGAGATCAG GAAGCGACTA TATTCCTTTA GGTCTCTATG
 253451 ATTCTCGAGA AGAAAAGTTA GTTTCTTTGG ATCTTCCTAT TACACGAGCT
 253501 GTAATTTTTG GTAATGACCA AGATTCGGCA AAGTCGTCAG A TACTGCGAA
 253551 CCACTATGTT TTATTTAATG ATTACATGCA GATTATTGTT TCTGAAGAGA
 253601 GTGGTTCTAT AGAAGGTATC AATTTACCTT TTGCTTCAAC AAATAATAAA
 253651 AGCATTGTGA ATGAAATTGG TTTTGATAGG GATTTAGCTT CAGAGAAATC
 253701 TCCTGAAGCT CTTTTCCCTG GGCTGTCTTC AAAACTTCCT GATGGCCAAC

253751 AAGCCAAAA CTCGATTGGA GGTACTACC CTTTATTGCG CAGGGGATTA
 253801 TTAAGTGATT CTAAGAAATT ACTTCCTCTA GAGTATCACG CATTAATGT
 253851 GGTTCAGGA AGAGAGCTAG CGACTCCTGT GGCTTTAAGA TACCGAGTTC
 253901 TTTCTATAC CCCCATTCC ATTCAATTGG AAAGCTTAGA TAGATCGGTT
 253951 CAGAAGGTAT ACAACTTCC AGAGAATCCG GAAGAAAAGC CCTATGTTTT
 254001 TGAAACTGCA ATTACTTTAA CGAAAGAAAC CGAAGATGTA TGGGTAACCT
 254051 CAGGAGTTCC TGAAGTGGAG ATCATGTCAA ATGCTTCAGC CCCAACCATT
 254101 AAATACAGGG TTATCAAAAA AAATAAGGGG TCTTTAGATA AAGTTAAGCT
 254151 TCCAAAAGTA AAAGAGCCTT TAGCTATACG TCGTGGTGTT TATCCTCAAT
 254201 GGATTTTAAA TTCGAATGGA TATTTCCGTA TTATTTTAAAC TCCGTTGTCT
 254251 GAAATTGCTT CTGGCTATGG ATCTCTCTAC ATTTCCGGTT CTACGGCTCC
 254301 GACAAGATTG TCTGCTATTT CTCCTAAAAA TCAACTGTAT CCAGTATCAA
 254351 AATATCCTGG ATATGAGACC TTGCTTCCTT TGCCAAAAGA TGCAGGGACA
 254401 CATCGATTTT TAGTGTATGC AGGTCCCTTG GCAGAGCCTA CACTTAAAGT
 254451 ATTAGATAAG ACAATTACTC AGGAGAAGGG AGAAAATCCT GAGTATCTTG
 254501 ATAGCATTTT TTTCCGTGGT GTTTTTGCAT TTATTACAGC TCCTTTTGCA
 254551 GCACTCCTAT TTATTATTAT GAAGTTCTTC AAATTGGTTA CGGTTCTTG
 254601 GGGAATTTCC ATTATTTTAC TTAAGTATT TTTGAAATTG CTTCTCTATC
 254651 CTTTAAATGC ATGGTCCATA CGATCTATGA GGCGTATGCA GATTTTATCT
 254701 CCTTATATTC AGCAAATTCA GCAAAAGTAT AAGAACGAAC CTAAGCGTGC
 254751 TCAGATGGAA ATCATGGGCT TGTATAAGAC AAACAAAGTG AATCCTATCA
 254801 CGGGTTGTTT ACCTTTATTG ATACAGCTTC CTTTCCTAAT TGCGATGTTT
 254851 GATTTATTAA AGTCATCATT CTTATTACGA GGAGCCTCGT TTATTCCTGG
 254901 GTGGATTGAT AACTTAACAG CTCCTGATGT GTTGTCTTCT TGGCAGACAT
 254951 CGATATGGTT TATTGGAAAT GAGTCCACT TACTTCCTAT TCTATTAGGT
 255001 ATAGTGATGT TCTTACAACA GAAGGTCACG AGTTTGCATA AGAAAGGACC
 255051 TGTTACGGAT CAGCAGAAAC AGCAACAAGT TATGGGGAAC ATGATGGCGA

255101 TTTTATTTAC CGCTATGTTC TATAACTTCC CTTCAGGATT AAACATCTAT
 255151 TGGCTTTTCGT CTATGATTTT AGGAGTCGTC CAGCAGTGA TACTAATAA
 255201 GATATTAGAT AGCAAACATC TAAAAATGA AGTGGTTTTA AATAATAAAA
 255251 AACATCGATA ATACTGAAAA AAAAGTCGAA TTGTCTTAAT ACAGGCTTAT
 255301 TGAAACAAGC GATTTATATG AGCTCATAGT TTATGCGAGC ATGGGAAGAA
 255351 TTTCTTTTGC TACAAGAGAA AGAAATTGGC ACAAATACTG TAGACAAGTG
 255401 GTTGCATCT TTAAGGTCT TATGTTTTGA TGCTTGTAA TTGTATCTTG
 255451 AAGCTCAAGA TTCTTTTCAA ATTACTTGGT TTGAGGAGCA TATAAGACAT
 255501 AAGGTTAAAT CTGGTCTTGT AAATAATAAC AATAAGCCCA TTCGTGTTCA
 255551 CGTTACTTCG GTAGATAAAG CAGCTCCTTT TTATAAGGAG AAGCAGATGC
 255601 AGCAAGAGAA GACAGCATAC TTTACCATGC ATTATGGAAG TGTGAATCCT
 255651 GAGATGACCT TCTCTAATTT TTTAGTTACC CCTGAAAATG ATCTTCCTTT
 255701 TCGTGTTTTA CAGGAATTTA CTAAGAGTCC TGATGAAAAC GGAGGAGTTA
 255751 CTTTTAATCC AATTTATCTG TTTGGACCTG AGGGATCTGG AAAAACTCAC
 255801 TTAATGCAGT CAGCTATCAG TGTTCTTCGT GAATCTGGAG GTAAGATTCT
 255851 CTATGTTTCT TCGGATTTGT TTACAGAGCA CTTAGTCTCT GCTATCCGTT
 255901 CAGGAGAAAT GCAAAAATTC CGTTCTTTTT ACCGCAATAT TGATGCTCTA
 255951 TTCATTGAGG ATATCGAGGT TTTTTCAGGA AAGTCGGCAA CTCAAGAAGA
 256001 GTTCTTCCAT ACGTTTAATT CTCTTCATTC TGAAGGGAAG TTGATTGTAG
 256051 TGTCTTCATC CTATGCGCCT GTGGATCTCG TTGCTGTTGA AGATAGATTG
 256101 ATCAGCAGGT TTGAATGGGG AGTTGCAATT CCGATACATC CTTTGGTTCA
 256151 GGAAGGATTG CGCAGTTTCT TAATGAGACA GGTAGAGCGC TTATCTATTC
 256201 GCATTCAAGA AACGGCCTTA GATTTTTTAA TTTATGCGCT ATCTTCCAAC
 256251 GTAAAGACCT TACTGCATGC ACTGAATCTT TTAGCAAAGA GGGTAATGTA
 256301 TAAAAAACTC TCTACCAAT TACTATATGA AGATGATGTG AAAACTCTTT
 256351 TAAAAGATGT TTTAGAAGCA GCAGGAAGCG TTCGTTTAAC TCCTTTAAAG
 256401 ATTATTCGTA ATGTTGCTCA ATATTATGGG GTCTCTCAGG AGAGTATTTT

256451 AGGACGTTCT CAGTCCCAG AATATGTATT GCCACGTCAG GTAGCCATGT
 256501 ACTTTTGTCG TCAGAAGCTT TCACTATCAT ACGTGAGAAT AGGCGATGTC
 256551 TTTTCAAGAG ATCATTTCGAC GGTAATCTCA TCCATACGAT TGATTGAACA
 256601 AAAAATAGAA GAAAATAGCC ATGATATTCA CATGGCTATT CAAGATATTT
 256651 CTAAGAATTT AAATTCCTTG CATAAGAGTT TGAATTTTTT CCCTAGCGAA
 256701 GAGATGATTA TTTAGAGGAG TAACGATCCT CTAATTTTTG CTTCTGGACA
 256751 GTTGCAGCAG CCGGCGCAGA TATCATTTTC ATTGAGTTCA ACTCCGCTGT
 256801 TCTAGCTTGT TTAAATTGTC TAATATGATT GAAAGCAAGG ACAAGAGCAA
 256851 TACCAGCAAG TACTATCTGT ATAGCAAGAA CAATGCCGAT TCCTAAGGAT
 256901 AGGGTGACTC CACTGGCAAT AACCAGTCCT AGGGTGGTTA AGGAAGCAAG
 256951 GAGGAGACCT AGACTAACTA AAATAGGCTT AAAAAATACC CAGGAACAAT
 257001 AACGATGTGT TGCCTTGTA GATACTGATG GTTGTGGTTG TGTAGTCTGA
 257051 GGTGTTTGTG CTACTGTAGC CATGTTGTTT TCCTGAGTAA AAAATTAAAT
 257101 AATTTTAATC TGTCTATATA ATTGTTGTCA ATAATATTTT TATAGATATT
 257151 ATTTTAATTT TATTAAAATT TATCAACTAA TTTTATTTAT AGTCTTTCTC
 257201 TAGTTCAGAA CTATTAAGCT CTCTTAATAG GCTCTCTTTT ATTTACCAAG
 257251 GCGAAGTTCT TTTTTTTTTT TCAAATTTTCG CATATGGAGG GTTCTAGATG
 257301 GTATTTGAAA TGGTTGCATT GTGGAAGATT TTTCGAGTTT TGATAAGAAC
 257351 AAAGTCAGTG TTGACTCTAT GAAACGGGCG ATTTTAGATC GTCTGTATTT
 257401 AAGTGTGTA CAATCACCAG AGTCCGCATC TCCTAGAGAT ATCTTCACAG
 257451 CTGTTGCAAA AACTGTTATG GAATGGTTGG CCAAGGGGTG GCTGAAAACT
 257501 CAAAATGGCT ACTATAAAAA TGATGTAAAA AGAGTTTATT ACCTTTCCAT
 257551 GGAATTTCTC TTAGGGAGAA GTCTAAAAAG CAATCTTTTG AATTAGGAA
 257601 TTCTAGATTT AGTAAGGAAG GCACTAAAAA CTTTAAATTA TGAATTTGAC
 257651 CACCTTGTA AAATGGAATC CGATGCAGGA TTAGGAAATG GTGGTTTGGG
 257701 GAGACTGGCA GCTTGTTACT TGGATTCTAT GGCTACATTA GCAGTTCCAG
 257751 CCTACGGCTA CGGTATACGC TATGATTATG GTATTTTTGA TCAGAGGATC

257801 GTCAACGGGT ATCAAGAGGA AGCTCCTGAC GAGTGGCTAC GTTATGGAAA
257851 TCCTTGGGAA ATCTGTAGGG GAGAGTACCT CTATCCCGTA CGATTTTATG
257901 GAAGGGTCAT TCATTATAACC GATTCTCGAG GGAAACAGGT GGCAGATCTT
257951 GTCGATACCC AAGAGGTATT GGCGATGGCT TATGATATTC CGATTCTCTGG
258001 GTACGGTAAT GATACTGTAA ATTCTCTAAG GCTATGGCAA GCACAATCTC
258051 CGCGAGGCTT TGAATTCAGC TATTTTAACC ACGGGAATA TATCCAGGCT
258101 ATAGAAGATA TCGCCTTGAT AGAAAACATC TCTCGCGTCC TCTATCCTAA
258151 TGATTCTATT ACTGAGGGGC AGGAATTGCG TCTCAAACAA GAGTATTTTT
258201 TAGTTTCAGC AACCATTCAA GATATTATCC GCAGATATAC AAAGACACAT
258251 ATTTGTTTGG ATAACCTTGC GGATAAAGTC GTAGTACAAT TAAACGATAC
258301 CCATCCCGCT TTAGGGATTG CTGAAATGAT GCATATTTTA GTCGATAGGG
258351 AAGAATTACC TTGGGATAAG GCTTGGGAGA TGAATACAGT CATCTTTAAC
258401 TATACCAATC ATACAATCCT CCCAGAGGCT TTAGAGAGAT GGCCTCTCGA
258451 TTTATTCTCT AAGTTATTAC CTCGGCATT TATGAAATAA
258501 ATTCCCCTTG GTTAGAAAAA GTTGGCTCTC GCTATCCTAA AAATGATGAT
258551 AAGCGCCGGT CTTTATCCAT TGTTGAAGAA GGGTATCAAA AGCGTATCAA
258601 TATGGCAAAC CTGCGCTAG TAGGTTCTGC AAAAGTAAAT GGAGTTTCGT
258651 CATTCCACTC TCAGCTGATT AAAGATACTC TCTTTAAAGA GTTTTATGAG
258701 TTTTCCCTG AGAAGTTTAT CAATGTGACC AATGGGGTGA CTCCACGACG
258751 ATGGATTGCT CTCTGTAATC CTCGTTTGAG TAAGCTTCTC AATGAAACTA
258801 TAGGGGATCG TTATATCATT GATCTTTCTC ATCTTTCATT GATCCGTTCC
258851 TTGCGCAAG ATAGTGGTTT CCGAGATCAT TGGAAAGGGG TAAAATTAAA
258901 AAATAAGCAG GATCTAACAA GTAGAATTTA TAATGAAGTT GGAGAAATAG
258951 TAGACCCTAA TTCTCTCTTT GACTGTCATA TTAAGCGTAT TCATGAGTAT
259001 AAACGACAAC TAATGAATAT TTCTTAGAGT CATCTATGTT TATAATGACT
259051 TGAAAGAAAA CCCTAATCAA GATGTCGTCC CTACAACAGT AATTTTTTCT
259101 GGTAAGGCGG CTCCTGGCTA TGTCATGGCC AAATCATT TCAAGTTAAT

259151 CAATAGCGTT GCTGACGTTG TAAATCAAGA TTCTCGAGTT AATGATAAGC
 259201 TTAAGGTTCT TTTTTTACCT AACTATCGAG TTTCTATGGC TGAGCATATC
 259251 ATTCCTGGTA CAGATCTTTC AGAACAGATT TCTACAGCTG GAATGGAGGC
 259301 TTCTGGAACA GGAAATATGA AATTTGCTTT GAATGGAGCT CTGACTATAG
 259351 GAACTATGGA CGGTGCAAAT ATAGAAATGG CAGAGCATAT TG GTAAGGAG
 259401 AATATGTTTA TTTTGGTCT TTTGGAGGAG CAAATTGTAC AACTGCGGAG
 259451 GGAATACTGT CCTCAGACAA TTTGTGATAA GAATCCTAAG ATCCGTCAGG
 259501 TTTTAGATTT GCTAGAACAG GGATTTTTC AATAGCAATGA TAAAGATCTG
 259551 TTTAAACCGA TAGTACATCG CCTACTGCAT GAAGGAGATC CCTTTTTTGT
 259601 CTTGGCTGAC TTGGAGTCTT ATATCGCTGC CCATGAAAAT GTGAACAAAC
 259651 TCTTTAAGGA ACCAGATTCA TGGACTAAGA TTTCTATTTA TAATACTGCA
 259701 GGAATGGGCT TTTTCTCTAG TGACAGAGCC ATTCAGGATT ATGCCAGAGA
 259751 TATTTGGCAT GTTCCTACAA AATCTTGCTC TGGAGAAGGA AATTAAGAAA
 259801 TAGACAGGAT GGAATCAGGG ACTCTTTCCA TAGCCAAGGA GCTATAGAAA
 259851 GAGTCCTTTT TGTTCAAAGA TTGCTAGTTT AATAGTAGGA CAGCCGGAGC
 259901 TTCTAAGATC TTTTGTAATC GTTTCATAAA CATCGCAGCA GGATAACCAT
 259951 CAATCACTCT ATGATCTACA GATAGGGTAA GATTGCAGGT AGATCCTATA
 260001 GTAATTTCTC CGTCAAGAAC AAGAGCTTGT TCTGTAACAC TTCCTACGGC
 260051 AAGAATCGCC GCTTGAGGAG GATTGACAAT CGCTGTAAAT TCAGTGATTC
 260101 CTGTCATTCC TAAGTTAGAG ACACAGAAGG ACCCTCCTTT GTATTCACTG
 260151 TCTTGAAGAG ATTGATTTCT TGCTTTTAAC GCTAAGCTCT TAATTTCTGC
 260201 TGAAATCATG CCGAGATTTT TACGGTCTGC GCAGCGTATA ATTGGCGTAA
 260251 TAATCCATC TGAATGGCC ACAGCTATCG AGATATCGAT AGTATCAAAA
 260301 CGGACGATTT TATTATCGAC ACTGTAAAT CCTGAATTGA TAGAAGGGAA
 260351 CTCTTTGAGC GCCAGAGCAC AGGCACGTAC AATGCAATCG TTAATAGAGA
 260401 GTTTGATTCC CTGAGCTTGA AGTTCTTTGA GCAGATTAAG GAGAGGTGAG
 260451 GCGTAGACCT GCTGCCTTAC ATAGAAGTGA GGAATAGAGA TCTTAGCAGC

260501 TTGTAGGCGT GCAGCAATCA CTTCCCGAAT CGGAGAGAGA TTCTCCTCAT
 260551 GATAGGAACC TGGAGGCACT TCGGGAGACT CAGGATAGCC AAAACCAGCA
 260601 ATGCTTTTAG GAGGAGCTTT CTCTAAATCT TTTTTTACTA TACGTCCTCC
 260651 AGGACCACTC CCTTGAATTG ATGAGACATC TATGTTTTTC TCTTTTGCTA
 260701 GTTGTCTAGC TAATGGAGAG AGATTATTCG TAGTGCCTAC GTGTTTGAAG
 260751 ACTAAAGGCG AGGAGAGAGG TGGCTCTGGC TTAAAAGTTA CTGCTGTGAA
 260801 TGTTGCTGAG GCAGCTTGTG GAGTTGTTGC AGGCGAGACC TCTTCAGAAG
 260851 AACCTTTTGG AGATGCTTCA AGGTTAGAAG GTTCTGTCTT AGGAAGAAGT
 260901 TCTTCTAGAT TAAAGGGCTC GTTGGCTTCT GTAGAGAGTA CCGCAATAGG
 260951 GGTGCCTATA ACGATTTTCT CGCCTTCATG ACGTAAGATT TCACGAATCC
 261001 AGCCATCTTC ATTTGCTGTA TGTTCTAAAA TAGCTTTGTC TGTAGAGATC
 261051 TCTACAATGA CGTCTCCAAA ACTGACCTGA TCATTACTTT TTTTATGCCA
 261101 TTCACTATA GTGCCCCACTT CCATAGTTGG AGAAAGCTTT GGCATTTTCA
 261151 ATAAGGAGAT CACAACTTA CCTCATGACT TTTTCAATGG TATCTAAGAT
 261201 TCGGTTAACA TTAGGCAAAG TGGCCTGTTC TAAGATTTTA CTATAGGGCA
 261251 TAGGCGTTTC TTTTGGCAT ACCCTTAAGG GGGGAGCATC AAGAGAATCA
 261301 AAAACATGCT CAGTAATCAG GGCAATAATT TCAGAAGAAA TCCCAGCGAA
 261351 GTAGTGGCCC TCTTCAATTA CAATACAGCG TGAAGTTTTT CGTACCGATG
 261401 ATAAAATTGT TGATATGTCT AAAGGTTTGA TCGTTCTTAG ATCAATAATT
 261451 TCTATAGACA AGCCCCAACG TTTTTTGGCT AGAGAACACG CTTCTTTTGT
 261501 AATGGAAACC ATACGGCTAT AAGTAATAAT TGTAAGGTCA TTTCTTCTT
 261551 GAACTCTATG TGCTTTCCCA ATAGGAACGA GATATTCTTC GGTGGGGACT
 261601 TCCCCTTTTA AGTTATATTC TAGCTCGTTT TCTAAAAAAA GAACGGGGTT
 261651 ATTATTTCTG ATTGCTGATT TTAATAAGCC TTTAGCGTCG TAAGGGTTCG
 261701 AAGGGGCTAT AATAATAAGA CCTGGAATAT TAGCATACAA CGACTCAACG
 261751 CAATGAGAAT GCTGGCAAGA TACCTGGGCT GCAGCACCAT TAGGGCCACG
 261801 AAAAATATA GGAACGGAAA ACTTCCCTCC AGTCATAAAA TGCATCTTAG

261851 CTGCATGAGA AATGATTTGG TCCAAGGCTA CAAAGGAAAA GTTCCAGCTC
 261901 ATAAATTCTA TAATAGGGCG CAGGCCTGAC AATGCGGCTC CTATTCCAAT
 261951 TCCAGAGAAG GCTGCTTCAC TAATAGGAGC ATCAATGACT CTCTTAGGGC
 262001 CCCATTTATC TAATAAGCCT TTGGTGACTT TATAAGCACC ATTGTAGTCA
 262051 CCAACCTCTT CACCAAGAAT ACAGACATTA GGATCGCGAG ACATCTCTTC
 262101 GTCAATTGCT TCTCGGAGAG CTTCTCGAAT TTCTAATGTT TTATGTTTAG
 262151 GCATAGACTC CTTCTCTAA TGTGGTGACG GATGGATCTG ATGAGAGTTT
 262201 TGC GTTAGAG AACGCTTCTA AAACAGCAGT TTTGCATTCT TGGCGTATAT
 262251 TTTGAAATTC CTCTTCAGTC AGAACCTCTA ATCGAATTAG CCAATCTTTA
 262301 GCTAGGACAA TAGGATCTTT TTAAATAAA CACTGCATTT CTTCTTTCGA
 262351 TCTATATAAA TTAGGATCTG ATATAGAATG CCCTCGAAAT CGGGAGCAGA
 262401 GACACTCAAC TAAAACCGGA GATTTCGGTAT CAACCATATA GCGATAAGCC
 262451 TCTCTAAATC CTAAAAGAGA GTTAAATAGA TCAAACCAT TGACTGTGAC
 262501 TGCACGGATA TCGTAGGAAC TTCCTTGAGA CTCTGCTATG GGCTGTTTTG
 262551 CAACAGCACG ATTTAATGAC GTTCCCATAC TCCAGCCGTT ATTTTCAATA
 262601 ATAAGCATTA GAGGGAGTTG GTGAAGAGAA ACAAAGTTCA GAGTTTCATG
 262651 GAATACACCT TGAGCTACCG CACCATCTCC GATAAAGCAT AGAGAACTC
 262701 TATTTTTTTG TTCTTGATAT TTGATGGTAA ATGCGGCTCC AGCTGCGAGG
 262751 GGAATTTGTC CTCCGACAAT ACCAAATCCT CCAGGGAAAT TAGGCCCACA
 262801 CATATGCATG GATCCTCCAC GACCTAAAGC GCATCCAGTT TCTTTCCCTA
 262851 AAAGTTCAGC AGCAATTTCT TGAAGGGGAA TGTTGAGAAG AATCGCAAGT
 262901 GCGTGGCAGC GGTATGAAGA GAACACCCAG GGATCTAGTC CTGTGTTTGC
 262951 GATTGCAGCA GTTGCTACAG CTTCTTGGCC AGCGTAAGAG TGGTAAAATC
 263001 CACCCACTAG CCCTTCTAGA TAGGCTTCTT CTCCTCGGGC TTCGAATTCA
 263051 CGAATCAGAA CCATCTGTTT TAAAAATTTA ATACAGGAAG CGGGCCCGTA
 263101 AAGGTCTAAG ATCCTTTCTA CTGTGGATTT CTCTGTGCCC TGAGAAGCTA
 263151 TATTATAAGG TGCTGAACTA TCCATAACTT TTTTATAAGG GAGACGCTTC

263201 GGGAGCGGTT TTTGATCTCA TTATAGGAGC CTAGGTTTCA AAAGACAAGG
 263251 GCTATCCGAA GAAAAAATAA TTAAGTTTTT TTAAAAATC GAAAAAATTT
 263301 TAATATGTTT CTCGTGGCAA TTTAAAGAGA AAAAAAGGA TTACATCACA
 263351 TGGACGTTTC TCGTAAAATT AATCGACACA CTCAGTTTTA CGTGGATTCTG
 263401 ATAGACGGTG TCATCAAAAA CTTTGATCAT AAGCCTAGTG AAGATAAATC
 263451 TCGAGATCAT GAAGAATTAG AGGAAAAGCT TTAACTATA AAAAAACGTA
 263501 TTGTTGCTTC TGCTCAAGAA TTTCAGAATC GTAAGACGGA CTCTAAAAAC
 263551 TACTACTTAA AAAAACTCA ATGGTTGCCA TTCAAAAATG AAGAGTTGGA
 263601 GCAAACCTAAG GAATTGTTTG CCATGTTAAC TTCGATGGAT AAAAAAATAG
 263651 CTCAGTTATT TTTTATTCT CCCGGATGTA GTAGCGACTG GGTAGAATTT
 263701 ACCGAAGTAA TTTGTCATTT AAATGATTCT ATAGGTTTGG GAGGGGTTTT
 263751 ACTGTGTTGC GGATTATTCG AACACAGTG TGAGCATGTT GTTACTGTGA
 263801 ATAAAAAGCT AGACCTACCG CTTCTTTTAG GAACTACCGT TGTAAATAGT
 263851 CTACGTTATT ATTTAACCTA TAGGAATATC TCTCTTTTGA ATTGTCAGAG
 263901 CATGAGTGAG CTAGGCAAGG AGCTCGGGGA TGTTTTAAAG CAACATGGAG
 263951 TTGCTTTCAC TTAAATTTTT AAAGAGATTG TGGATATAGA TTTATTAAAC
 264001 TATGTGAAAC TAATTCAAGG ACTCAAGCGT AGTGGAATA TTCAAGCAGC
 264051 TATTTATGAT AATGATGTGC CTACTTTACC TTCAGTTTCA TCGAGTCCTA
 264101 TTGCGTTGAG ATATAGCTTA GCGAATACAA TTCGGGGCCT TGCTTTACAT
 264151 GTAGATTTTT CTTCATTGAA GTTTATAAGC CCGTCTATAC TTTCCAATAC
 264201 AGAGCACACT GCAAAAGCTT TAAACTCTGG CGGAGAATGT TTTATCTTTT
 264251 CTAATCTGGA TGAATTCAAT CTTGGAATGA AAATAGTCAT GCAACTATTG
 264301 CGGACAGGGA AGATTAGCCC AGAAATTTTA AATAAAAATA TCATGAAAAT
 264351 TCTCATGATA AAAAGAAGAG TACGTTCTTT ATATATTTAA TTATTGGTC
 264401 TTTTTTTTGA TTTCAAACT AATAAAATAA ATAGAATTTT AGATCTCTGA
 264451 AGGAATCTCA GCAAGGCTGG GAGTCGATAG ATCTCTTACT TGTTTTCTA
 264501 ACTTACTTAG TCTTCTTCA GTTTTAGGAA GGTTCCGAAT TTTAGCAATC

264551 AACCGATGTG TTTCTTGATA AGGTCGTGCT GGAGCGCCTC CATAAATGCC
 264601 TGGAGAGGTG ATAGATTTTG TGA CTCCAGT TTGAGCAATC ATGATCACAT
 264651 GGTCTGCAAT AGAAATATGC CCAGTAATTC CGGTTTGCCC TCCAATGATG
 264701 ACATGTTTAC CAATTTTTGT AGAACCTGCA ATGCCTGCTT GGGCAACAAT
 264751 AATACTATGC TTTCCAATTT CTACGTGATG AGCTACTTGT ACTTGGTTAT
 264801 CTATTTTAGT TCCTTCATGG ATCACGGTGT TCTTGAATCG ACCACGATCT
 264851 ATCGTAGTGT TGGCTCCGAT TTCTACATCA TCACCTACAA TCACATAGCC
 264901 TAGATGCTTT AAAGGTTTGT GATGACCAA AGCATTTGTA ATATAACCAA
 264951 AACCACAGGA TCCTAAAACA GCTCCAGGTT GAACAAC TAC ACGGTTTCCC
 265001 ATGAGGACTC TTTCTCGAAT CACCACCTTA GGGTGAATCA GACAGTTAGC
 265051 ACCTAGAACG CTGTGAGCTC CAATGACACT TCCAGCTCCG ATGTATGTGT
 265101 CAGAGCCGAT ATGGGCATGT TGA CTAATGA CAACGTAAGG TTCTATGGTT
 265151 ACATTTTTCT CAATACGTGC AGTAGGATGA ATCACTGCAG TAGGATGAAT
 265201 ACCAGGAAAC CCTGATGTTA CGGGTTCAAT AAACAAC TCT ATGCACTTTT
 265251 GAAATGTTAG AGAAGGGGAT TCATTGGTAA TAAGAAAGTT TTTCTTTAGG
 265301 TGGGCATGTT GCATTGCCTG AGATCTAGAT AAAATAATAG CACCAGCTTT
 265351 GGTGTTTTTT AGAAAGCTAG AGTATTTCTC ATTATCTAAA AAAGCAATAT
 265401 GGTGAGGTTG CGCCTGACTA ATATCTTCAA CACCTGAAAT AGGAGTTTCT
 265451 ATATTTCTT GAAC TCGAC TTGTAGTAGC TCAGCTAACT GTTTAAGAGT
 265501 GTAGACTGGT GCTTCGGACA TAGAAAAC TCT CTTAAACTTG GACTAGTTTT
 265551 GTTTTTTGAA AGATTCGTTA AGAATAGCAA TAATTTCCGT TGTTTTATCA
 265601 GTCCAGGTG CTATTGCTAA GACAGCTTCT TCATTAAGGA TAGCTTCTAG
 265651 TTTTCTTTG GACCGCACTG ATTCTGCAGC TATTTTACT TCTTGAATGA
 265701 GTTTTTGAAT GCGTTTTACA TG TACTTTGA TTGATAGATT GATAGTACTG
 265751 AGACTGGTAC GCATTGTACT CTCCTGAAAG ATCTTCGAAT TTCTTTCGCA
 265801 ACTCTTCAGA GGCAGAATCC GATAGGCTTT CCATGTAATC TTCATCTTGC
 265851 AACTTATTAT AAATAGAAGT GAGTTCTTCT TCTATTTTCT CAGCATTTTT

265901 TACAAACTGC TGTTTCATAG CTTCCAATTC TTCAGTTTCC TTTTACCTA
 265951 GATCGGATTC TTCAAGACAT CGCTTTAAAT TAACATAGCC TAAATTTGCA
 266001 TGAGCTGCGC TTGTTGATCC TAAAACAAGA AGAAATGTAG AAAATAATAA
 266051 TTTTTTCATA ATTACCTGTA CTAACCTGGG TATAGAAAAA GGGTACCAAA
 266101 AGCCTTTTCT GAAAACAACA AAGATTTCTT TCGATAAGTC CTTAATTTAT
 266151 ATCTTAGAAC ATGCCCCCTA AAGCAAAGAA GAATCGCTGA GATACATCAA
 266201 TTTTTTCTCC ATTCAAAGTC TCGGTTGGAC GGAAGGGCCA ACCAAATCCT
 266251 AACATAACAG GAACATTATT CATTACATCG AAGCGCAGAC CAAATCCAGC
 266301 ACTACTACGT AGATCTTTTA ACGAAATCTT ATACTCTTGT AAACCGACAA
 266351 AACCTGAGTC TAAGAATACA AAGGCACTAA TATTAGGTTG TCTGATGAGA
 266401 GGGTATTGAA ACTCTTCTGA AATAAGGAGC GAAGAGAGTC CTCCCTGAGG
 266451 TTCTGTAGCA GAGTATTTTG GACCGATAAT AAAGGATTTA TATCCCCGAA
 266501 CTGTAGTCTC TCCACCTAGG AAGAAGCGCT CACTGACAGG AACTCCTTCA
 266551 GCTGTAGTAT TGCTATAGGG TTTAATAAAT TGAGCTTCCC CTTTGATTTT
 266601 CAAAATACCT TTACGCGTAA GTTTTCTATA GATAGAGCTG TTTAAAGAGA
 266651 GTTTTGTAAG ATGATAAGTT CCTCCCAAAC CAGAAACCTC AAAAGTCACC
 266701 CCCCCGCGAA TCCCTGTAGT TGGAGTTCTA GGACTATCTA CAGAATCGTA
 266751 ATTCAAGTTG ACACCTGCAG CAGAGACAAA TCCTTTATTG CTGTCTATAT
 266801 TTGGCCCTAG GAGGAACTTA CGTTTTTCAT GTAAACTCGT TTGACTTCCT
 266851 CGATAAAATA GACCGTATTT CAGGTGTTCTG TTCAAGATAT ACGTTGTGCT
 266901 GACGTTCCCG CCATAGGTTT GGACAGCATA ATCTTTAGAT AATGCTCTGT
 266951 TAATTGATTT ATCTAATTCA ATTCCTAAAA TCCAAGGAGT GTTTAGAAAA
 267001 TGAGGTTTGG TCCACTTCAA AGTATAGTCT GTGACTTTGT CCCCAGGTT
 267051 GGCTTTTAAG AATAGATGTT CTCCACCGCC TCTTAGACAA CGAAAACCTT
 267101 TAGAAAATAT ATTTCTAGCT CCAAATAGAT CAAAATTACT TTCAGATAGT
 267151 TCAATTCCTC CAAAAGATT GTCAAGAGAA CTAAATCCTA AGAATAAGCC
 267201 TAAGTTTCCT GTTGTTGTTT CTTTGACTTC TACAAAATA TCTCGGTATT

267251 GATCCGCATT GCCCATAGGA TCAAGTTGAG AACGAACTGT ATAGACACTA
 267301 ACGCTTTGGA AGTAGCCTGT ATTTCTTAAA CGTTGCTCAG TATCTTCTAG
 267351 CTTTAAGCGA TTGAATGTAT CTCCTGGGAA GAGACTGGTT TCGTGTAATA
 267401 TAACGTCAGA TTTTGTATGG GTATTCCCAG TAATTTTAAT TAACCCAACT
 267451 TTATAAGGAG ACCCTTCACT TACCTCATAA GTTACATCAT AAATAGGGCG
 267501 GGTTCGCTGA GGGATGAAGA GAACGTCTAC ATTGGTATTG ATGTAGCCAT
 267551 ACTTTGCATA AGTTTGTTTG ATCTTATGAG CCCCATCCCA TATTTTATCG
 267601 GGGCAATAAA GATCATTGGG GCCGACTTGG GATTGCTTTT CTATAAGGCG
 267651 TTTTGGCAAA ACCTCAAACC CTTGGATATG GACGTGTCCT AAGGTATATC
 267701 GCGACCCTCG ATCAATATCC ATGTAAAGAA GAATATTCCC TTTGTCGTCA
 267751 AGGTCATAGT GAGAGTTGAC TATAGCATCA GCGTACCCGT TATTATGTAG
 267801 GTAATTCGTA ATTGCCAAGC TATCTTGTTT AACAATATCT GGGTGATAGA
 267851 GTCCAGCTCC AGTAAACCAA CTTGTAGTTG TAGAGTGCTG CTTGGTTTGA
 267901 ATAAATTCTT GGATATCTGA TTTTCTGAT CGAGAGATTC CTGAGAACGT
 267951 AAGCTGTTTA ATTTTCCCGC AAGGACCTTC ATTGATTTTA ATTAAAACAT
 268001 CGATGTGACC TTTTCTTGA TTGTGTTCCA GACTGTAGTC TACACTGGAT
 268051 GCGAAATATC CTCGCTTGAG ATAATACGTT CTTAGATCAT CAAGACCCTT
 268101 AAGAAATTTT TCTCGTTCAA AGAGATCATT ACGGTAAATT TGTAGGGTTT
 268151 TAAGAATTTT ATG TTCAGGA ACGACTTGAT TTCCTGAGAT ATGAATATTT
 268201 CGAATTGAGG GTTTAGCTAT TAGGTGAAGG GCTATGTTAG TTTTCCCTTC
 268251 AGAAAATTCT ACTTTAGGCT CAACAGAGTC GTATTCTTTA GCTAGAATTC
 268301 TCAAGTCTTC ATCAAAATCT AATTGAGAAA AAAGAGCCCC ACTTCTGGTC
 268351 TTTAATTTGG GTAAGGGATG TTTATTTGAA GCATTTTCTC CTTCGGTTAT
 268401 GATTGTGATA GAGTCTACCA CCACATGGCC TTCTTTAACT TTTTCAGTAG
 268451 AAAATAAAGT TAAAGGGGTT TGGATTAACG CTAGAATAGA TATTTGCAAG
 268501 ATAACTTTAT TTCGCATGAT GAGCATTCCC AAGAGTCTTC CCTAGATAGA
 268551 AAGCTTGTTT ATCAGATAAT GAGGTAAGCA CAAAATAGG GACAAAGAAG

268601 TTTTAAACA AGACCTTACT ACTTAGATTA TTTTGTATT TCTAAAGAAA
 268651 TTAAATTGT GTTTGATTTT TTCCCGTCCT AACACCGGGA TTATTACTGT
 268701 GCTCAGACTA GAGCCTGAAT ATTTCCGGCGG GTTTTTTGCT TCCTTAAAAA
 268751 GCTTAATAAA CCTTGTGCT CAAATAACTC TACTGATTAC ATGGTTTTAA
 268801 GGCAAATTCG ATAATTCACT GTGTCCTAGG GGAAAGTAAG AAAGACTAGC
 268851 ATAGAAGAGA ACCTTGTTTA CTTTtaggat AAGAGAGCTG CTAATAGGAG
 268901 TGTCGTCCAG AAAAAGCTCT TGCCAGTGTC CCTGAATCTA CATAATCAAA
 268951 AGATAAGCCT ATAGGAAGAC CTAAAGCTAG ACGGGAAATA TTTACAGAGA
 269001 AATGTTGTAA TTCTTGTTTT AGAAAAAGGG CAGTAGCATC TCCTTCTAAG
 269051 GTTGCATCAA TGGCTAGGAT AATTTCTTTT GGGCATAGCG TTTCTATGCG
 269101 TGATTTTAAA ATGGAGAGAC GTCGTTTTT TATATGTTTC CCTGTAATGG
 269151 GCGATAAGAG TGAACCAAGA ACATGATAAC GTCCCTTGAA TACTTTAGAA
 269201 CGTTCTAGAA AGAAAACATC TTTTGGAGAA GCGACAATAC ATAGACTTTG
 269251 GTTATCTCTT TCTTCTCTAC AAAAGTGACA GTCTGCCTCT TTAGATTCTT
 269301 TGAGAGTAAA ACATAGGGGA CAGTGACTAC GCTCACTAGC AACATTATGA
 269351 AAAGCGTTAC CTAATATTTT TAATTGTTTC CTGTCCCAAG AGATGAGTTC
 269401 AAAAGCAAGT TTTTCTGCTG TTTTAAATCC AATTCCTGGA AGTTTTCGTA
 269451 AAAAGAAAAT TAATTTAGAT AAGTAATCTG GATATCTTGT CATAGTGATG
 269501 TGAATTTTAT TTTTACATTC GGGTCTGGGG CCTAAATTAA GGTTAGAGTA
 269551 TAAACTCTCT GAATAGTATA CTAGCTTTTT TCTTTACATG TGGTCTCTG
 269601 TGAATAAAAA CAAAAAGCA GCAATTTGGG CAACGGGTTC CTATTTGCCT
 269651 GAGAAAGTTC TTTCAAACGC AGATTTAGAA AAAATGGTAG ATACCTCTGA
 269701 TGAGTGATC GTGACCAGAA CGGGGATCAA AGAGCGTCGT ATTGCTGGAC
 269751 CTCAGGAGTA CACTTCTCTT ATGGGAGCCA TCGCTGCAGA GAAAGCTATA
 269801 GCAAATGCGG GTTTAAGCAA GGATCAGATT GACTGTATCA TTTTCTCGAC
 269851 AGCAGACCA GATTATATTT TCCCATCAAG CGGAGCTCTT GCTCAAGCAC
 269901 ATTTAGGCAT TGAGGATGTC CCTACATTTG ATTGCCAGGC GGCTTGTA CT

269951 GGGTATTTGT ATGGTTTGTG TGTAGCTAAG GCTTATGTAG AATCAGGTAC
 270001 ATATAACCAT GTATTGTTAA TTGCTGCTGA TAAGTTGTCT TCTTTTGTAG
 270051 ATTATACAGA TCGGAATACC TGTGTGTTGT TTGGAGATGG AGGAGCTGCT
 270101 TGTGTCATAG GGGAGAGTCG GCCAGGATCT TTAGAGATTA ATAGGTTGTC
 270151 TTTAGGCGCA GATGGTAAGC TAGGAGAGTT ATTAAGCCTT CCTGCTGGAG
 270201 GTAGTCGTTG TCCTGCTTCT AAAGAGACTT TACAATCAGG CAAACATTTT
 270251 ATTGCTATGG AGGGAAAAGA AGTTTTTAAG CATGCTGTGA GACGTATGGA
 270301 AACGGCAGCT AAACATTCTA TAGCCCTGGC AGGCATTCTAG GAAGAGGATA
 270351 TAGATTGGTT TGTACCTCAT CAAGCTAATG AAAGAATAAT AGATGCTTTA
 270401 GCGAAGCGTT TTGAGATTGA TGAGTCTAGA GTGTTTAAGA GTGTACATAA
 270451 GTATGGAAAT ACTGCGGCCT CGTCTGTGGG CATTGCTTTG GATGAATTAG
 270501 TTCATACAGA ATCCATTAAG CTTGATGATT ATTTACTTTT AGTTGCCTTT
 270551 GGGGGCGGTT TGTCTTGGGG CGCAGTAGTT TTAAAGCAGG TCTAATAAGG
 270601 ACGATAATTT CATGAAAAAA CGTTATGCTT TTTTGTTCCC AGGACAAGGG
 270651 AGCCAATATG TAGGTATGGG ACAAGACCTA TATATGGAGT ATCCTGAGGT
 270701 TAGAGAGCTT TTTGATTTTG CTAATGAAAG GTTAGGATTT TCTCTGACTT
 270751 CAATTATGTT TGAAGGTCCT GAGGATCTTT TGATGGAAAC AGTACATAGT
 270801 CAGCTAGCTA TTTATCTTCA TAGCATGGCT GTGGTAAAGG TTCTATCTCA
 270851 GCGTTCTTCT ATTCAGCCTT CTTTAGTCTC TGGATTAAGT TTAGGGGAGT
 270901 ATACTGCTTT AGTTGCTTCC GATAGAATCT CCGTGCTCGA CGGCCTTGAG
 270951 CTTGTTAGAA AGCGTGGTCA GTTAATGAAT GAAGCTTGTA ATCAGAGCCC
 271001 AGGGGCTATG GCGGCTTTAT TAGGGCTTCC CTCTGAAGTT ATAGAGGAAA
 271051 ATATAACAAG TCTTGGTCAA GGAATTTGGA TTGCTAATTA TAATGCACCC
 271101 AAACAGCTTG TAGTGGCTGG AATAGCAGAA AAAGTAGACC AAGCGATTGA
 271151 GTTATTTTCGT GATTTAGGAT GTAAAAAAGC AGTTCGTTTA AAGGTGTCTG
 271201 GAGCATTTCA TACTCCTTTA ATGCAAGTTG CTCAAGATGG CTTAGCTCCA
 271251 GACATTTATG CTTTATGCAT GAAAGATTCT AGCCTTCCCT TAGTGTACACA

271301 CGTGGTAGGA AAATCTTTAG TAAATACTGA AGAAATGCGA GAGTGTTTAG
 271351 CTCGGCAAAT GACATCACCT ACGTTATGGT ATCAGAGTTG TTACCATATC
 271401 GAATCAGAGG TGGATGAGTT TTTAGAATTA GGTCCAGGAA AAGTTTTGGC
 271451 TGGTTTAAAT CGCTCTATAG GGATTTCTAA ACCGATTACA AGTCTTGGTA
 271501 CTTTTGCTCA GATTGAAAAA TTCCTATCAG AGGTATGATT TGTATGGATA
 271551 TAACATTAGT AGGCAAAAAA GTTATAGTAA CTGGAGGATC TCGAGGAATT
 271601 GGA CTCGGA TAGTTAAGCT TTTTCTTGAG AACGGAGCAG ATGTAGAAAT
 271651 TTGGGGATTG AATGAGGAGC GAGGTCAGGC TGTATAGAA AGTTTAACAG
 271701 GCTTGGGTGG CGAAGTTTCT TTTGCTCGTG TGGATGTGAG TCATAATGGT
 271751 GGAGTGAAAG ATTGCGTGCA GAAATTTTGA GATAAGCACA ACAAATAGA
 271801 TATTTTGGTA AATAATGCAG GCATTACCAG GGATAATTG TTGATGCGTA
 271851 TGTCTGAGGA CGACTGGCAA TCGGTGATTA GCACCAACTT GACTTCCTTG
 271901 TATTATACAT GTTCCTCAGT GATTGCGCAT ATGATTAAGG CGCGTTCAGG
 271951 ATCTATTATA AATGTGGCTT CTATTGTTGC TAAGATCGGT AGTGCGGGCC
 272001 AGACCAACTA TGCTGCTGCT AAAGCTGGGA TTATTGCTTT CACAAAATCT
 272051 TTAGCTAAGG AAGTAGCTGC AAGAAATATT CGTGTCAACT GCCTTGCTCC
 272101 AGGCTTTATT GAAACAGACA TGACAAGCGT GTTGAATGAC AATTTAAAAG
 272151 CTGAGTGGCT TAAGTCGATC CCTTTAGGTA GGGCTGGCAC TCCAGAAGAT
 272201 GTTGCTCGTG TGGCGTTGTT TTTAGCCTCG CAGTTATCGA GCTATATGAC
 272251 CGCGCAGACA CTGGTTGTTG ATGGGGGATT GACTTACTAA GACAATAGAA
 272301 GAAAGGGATT TGAAAATTTC TCTTCGAGAA CTAATTAAGT AACCGTCGAA
 272351 TAAAAAATGA TTTTTTGCGA TACTAATTCT CTTTCTCTTT GTCCCTAGGG
 272401 AATAGTGAAG TATTGTATAG TTAAATAGT AAAAGGATAT AAGCAATGAG
 272451 TTTAGAAGAT GATGTAATAG CAATTATTGT TGAGCAGTTA GGAGTGGATC
 272501 CAAAAGAAGT TAATGAGAAC TCTTCTTTTA TTGAAGACTT GAATGCTGAT
 272551 AGTTTAGATT TAACAGAATT GATTATGACT TTAGAAGAAA AATTGCTTTT
 272601 TGAAATTTCA GAAGAAGATG CTGAGAAGCT TCGTACTGTC GGGGATGTAT

272651 TTACTTATAT TAAGAAACGT CAAGCTGAAC AATAAACTTT CTTATATTCT
 272701 GGGTGCCTCG TAATCTTTTT TAAAGCTATT GTTTTTCAAT AGTGATTACG
 272751 GCGTGCCTTT TTTTCTATAG AGTAGCTCAT CTAGAAAGAT TTATTTTGTC
 272801 TTTTAAAGGT TCTCTGAACT TGATTTGTTT AGCATAGAGC TCTAAAAAG
 272851 ATAAAGCTAC GGATGGGCAC TCTCCACAA TGTTTAGAAT TTGTCCTTTG
 272901 CTAAGAACTA GCATGCGGAC TTGTGTATTT GCAGAAGCAT TGTATTCCCT
 272951 GGGCTTATTA TTGAATAAGC TTTCTCTCC AAAACAATCT AAAGGTTTTA
 273001 AATTTAGAGG AGACTCTAGT TTTTCTTTAG AGATCGTAAT GTATCCTTCT
 273051 ACAATGATAT AAAAGCTGAA TCCAGGTGT CCTATAGAGA ATACATTGCT
 273101 GCCAGGCTTA AATATTATCG TTTCAGTTTT ATCGGCAATT GTTAAAAGAA
 273151 GGTCCATGTC TAAAGATTGG AATATAATCG TTTTTTTTAG TAGAAAGGCG
 273201 CGATCGATCA AATTCATAAA AAAGTTCCTT ATTCACACCA TAGTTTtagT
 273251 TTTT